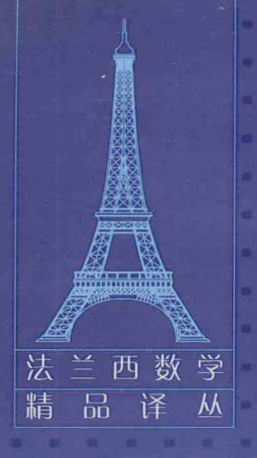


"十一五"国家重点图书

代数学教程

□ R. 戈德门特 著
□ 王耀东 译 张小萍 校



"十一五"国家重点图书



代数学教程

□ R. 戈德门特 著

□ 王耀东 译　　张小萍 校

DAISHUXUE JIAOCHENG



图字：01-2008-2672 号

图书在版编目(CIP)数据

代数学教程 / (法) 戈德门特著；王耀东译. — 北京：高等教育出版社，2013.6
(法兰西数学精品译丛)
ISBN 978-7-04-028757-8

Ⅰ.①代… Ⅱ.①戈… ②王… Ⅲ.①代数－教材 Ⅳ.①O15

中国版本图书馆CIP数据核字(2013)第068918号

策划编辑 王丽萍　责任编辑 李华英　封面设计 张 楠　版式设计 余 杨
责任校对 杨雪莲　责任印制 韩 刚

出版发行	高等教育出版社	咨询电话	400-810-0598
社　　址	北京市西城区德外大街4号	网　　址	http://www.hep.edu.cn
邮政编码	100120		http://www.hep.com.cn
印　　刷	涿州市星河印刷有限公司	网上订购	http://www.landraco.com
开　　本	787mm × 1092mm　1/16		http://www.landraco.com.cn
印　　张	37.75	版　　次	2013年6月第1版
字　　数	750千字	印　　次	2013年6月第1次印刷
购书热线	010-58581118	定　　价	89.00元

物 料 号　28757-00

《法兰西数学精品译丛》编委会

《法兰西数学精品译丛》序

随着解析几何及微积分的发明而兴起的现代数学, 在其发展过程中, 一批卓越的法国数学家发挥了杰出的作用, 作出了奠基性的贡献. 他们像灿烂的星斗发射着耀眼的光辉, 在现代数学史上占据着不可替代的地位, 在大学教科书、各种专著及种种数学史著作中都频繁地出现着他们的英名. 在他们当中, 包括笛卡儿、费马、帕斯卡、达朗贝尔、拉格朗日、蒙日、拉普拉斯、勒让德、傅里叶、泊松、柯西、刘维尔、伽罗华、庞加莱、嘉当、勒贝格、魏伊、勒雷、施瓦茨及利翁斯等这些耳熟能详的名字, 也包括一些现今仍然健在并继续作出重要贡献的著名数学家. 由于他们的出色成就和深远影响, 法国的数学不仅具有深厚的根基和领先的水平, 而且具有优秀的传统和独特的风格, 一直在国际数学界享有盛誉.

我国的现代数学, 在 20 世纪初通过学习西方及日本才开始起步, 并在艰难曲折中发展与成长, 终能在 2002 年成功地在北京举办了国际数学家大会, 在一个世纪的时间中基本上跟上了西方历经四个多世纪的现代数学发展的步伐, 实现了跨越式的发展. 这一巨大的成功, 根源于好几代数学家持续不断的艰苦奋斗, 根源于我们国家综合国力不断提高所提供的有力支撑, 根源于改革开放国策所带来的强大推动, 也根源于很多国际数学界同仁的长期鼓励、支持与帮助. 在这当中, 法兰西数学精品长期以来对我国数学界所起的积极影响, 法兰西数学的深厚根基、无比活力和优秀传统对我国数学家所起的不可低估的潜移默化作用, 无疑也是一个不容忽视的因素. 足以证明这一点的是: 在我国的数学家中, 有不少就曾经留学法国, 直接受到法国数学家的栽培和法兰西数学传统和风格的熏陶与感召, 而更多的人也或多或少地通过汲取法国数学精品的营养而逐步走向了自己的成熟与辉煌.

由于语言方面的障碍, 用法文出版的优秀数学著作在我国的传播受到了较大的

限制. 根据一些数学工作者的建议, 并取得了部分法国著名数学家的热情支持, 高等教育出版社决定出版《法兰西数学精品译丛》, 将法国的一些享有盛誉并有着重要作用与影响的数学经典以及颇具特色的大学与研究生数学教材及教学参考书, 有选择地从法文原文分批翻译出版. 这一工作得到了国家自然科学基金委员会数学天元基金的支持和赞助, 对帮助并推动我国读者更好地学习和了解法国的优秀数学传统和杰出数学成就, 进一步提升我国数学 (包括纯粹数学与应用数学) 的教学与研究工作的水平, 将是意义重大并影响深远的, 特为之序.

李大潜

2008 年 5 月

序言

在科学系的数学教学计划中, 随着相对现代和扩展的代数概念的引入, 在法国编纂一本便于初学者参考的著作变得刻不容缓, 呈现在读者面前的这本书尝试满足这个需要. 以下我们来概述指导撰写本书的主要构思.

首先, 本书以本人在巴黎讲授的在一般数学证书框架内的课程为基础, 因此阅读本书除了在中学教育中获得的知识, 不再要求其他知识. 其次, 我们补充了一些必要的几乎覆盖学士学位代数教学计划的知识 (这涉及目前的数学 I 证书), 该计划与一般数学的差别不大. 于是本书应该可以在许多年内被纯粹和应用数学的大学生使用, 它也几乎满足了对于具有一定程度理论性的物理感兴趣的人们的需要. 掌握了本书内容之后, 读者可以直接接触真正严谨的专门著述, 不过我们显然并不希望普通读者在进入科学系一年之后立刻这么做!

本书的主题是那些如今所有人都认可的对于将来的数学家和物理学家不可或缺的内容: 集合和函数, 群, 环, 域, 复数; 向量空间, 线性映射, 矩阵; 有限维向量空间, 线性方程组, 行列式, Cramer 公式; 多项式, 有理分式, 代数方程; 矩阵的化简. 这些主题的选择显然反映了过去 50 年内数学的发展, 但是我们认为这个发展应当以现今在面向专业数学家的著作中所保持的风格来表述.

许多人, 尤其是大多数局限于数学**应用**的人, 认为当为初学者写书时, 试图做过分严格的证明, 个个定理皆证, 引进太一般的概念, 使用严格定义和干巴巴的术语, 是无用的, 甚至是有害的. 如果他们果真有理, 那么这似乎说明, 与有见识的专业数学家的观点**相反**, 数学课本编写得越不严密, 初学者反而越容易理解. 职业拉丁语学者, 作为他们的工作, 懂得每天从意大利地下挖掘出的记录片段, 但是没有一个拉丁语教授会利用这些记录教初学者拉丁语 —— 人们更愿意借助精心编写的语法进行

拉丁语教学. 在数学中也一样, 当遇到解释一个艰涩的定义, 补充一个不完善的证明, 或者揭示一个定理的内在理由这些事情时, 人们自然不能奢望初学者和专业数学家有同样的感受.

还必须注意到, 20 世纪初数学的进步使有诚意的人们进行数学教育实质性的变革成为可能. 一些简单且普遍的新概念显著地扩展了传统推理的应用范围, 一些新的证明发现后, 使得昔日对于初学者太困难的证明, 现在也可以接受了. 而且对于严格性的关注, 原先总认为是数论专家们的特长, 30 年来已渗透到数学的所有分支, 现在还 (不同程度成功地) 渗透到部分学校教材, 有些人还明显地走在了专业数学家的前面. 这种变革以及有时与之伴随的对于变革成果的夸大宣传, 并没有阻止一些数学应用者的抗议, 他们因难于理解其子女的教材而烦恼; 有时还会听到人们对于数学家的指责, 说他们夸大自己贡献的重要性, 并且转移初学者对于更实际的问题的注意力. 无疑这类指责不无道理; 但是人们为什么不议论一下, 比如说, "探索宇宙" 的专家, 他们总觉得请求巨额资金去认识金星是天经地义的, 可在他们的眼皮底下, 却还有几百万人在为免于饿死而苦苦挣扎. 数学至少有一个好处, 那就是花费不多.

面对一些人当中激起的日益严重的憎恶和惊慌情绪 —— 这种情绪在 Paolo Ucello(*) 的《亵渎圣饼》中如此淋漓尽致地表现出来 —— 必须挑明我们同众多人士之间存在的意见分歧. 他们要求一般科学家, 尤其是数学家, 要培养大量为人类生存急需的技术人员. 而我们认为在我们生活于其中的、科学上和技术上超发展的 "伟大" 国家里, 数学家乃至许多其他人的首要任务是培养他们并不需要的人 —— 即能够独立思考, 摆脱错误推理和模棱两可话语的人, 这些人认为真理的传播, 举例说, 比彩色的和立体的卫星电视更重要, 他们是自由的人而非技术统治下的机器人. 遗憾的是, 培养我们所缺少的这样的人的最好方式是不给他们讲授数学科学和物理学, 因为知识的这些分支是不管大量人类问题的, 高度文明化的社会的首要任务是调和这些人类问题, 但这是应当受到质疑的. 即使如此, 你至少可以尝试在数学教学中, 激发人们对于自由和批评的欲望, 并且使他们习惯于把自己看作具有理解能力的人.

为了适合作为本书对象的初学者, 我们努力以专业数学家的语言与他们交谈, 明确地并且一劳永逸地定义所有数学术语, 清晰地陈述所有定理, 并且除了为使本书保持合理的篇幅的少数例外, 完整地证明这些定理(**).

还有, 我们在建立形式尽可能一般化的定理时努力遵循以下显然的原则, 即对于初学者, 坚持一般化不能导致简单结果的证明变得本质上复杂了, 也不能使一般

(*) Paolo Ucello (保罗 · 乌切洛, 1397 — 1475), 原名保罗 · 迪 · 多诺 (Paolo di Dono), 意大利画家, 以其艺术透视之开创性闻名. 由于他所绘飞禽精致而有 "飞鸟" 之称, "乌切洛" 即意大利语飞鸟之意, 其著名作品包括描述 "圣罗马诺之战" 的三件套油画. —— 译者注

(**) 大部分没有证明的命题在 §0 到 §5. 对于初学者, 叙述集合论和形式逻辑, 而不承认许多 "显然" 的结果, 这不成问题. 关于逻辑推理的 §0 的目的不仅是要向初学者指出 "许可的" 推理和 "不许可的" 推理 (这对于试卷的阅读是必不可少的), 而且告诉他们 "数学哲学" 并不必然归结为没有结构的文字游戏.

化的结果实际上是没有用的. 在线性代数中就是这样 —— 人们通常把问题归结为有限维实向量空间 —— 我们总是至少采用任意交换域上的向量空间的观点, 或者甚至非交换域上的, 那里取消了交换性假设; 仅用到加法和乘法的最简单的概念, 甚至是对于任意环上的模表述的. 模在分析之外的数学的所有其他分支里, 都扮演至少与向量空间一样重要的角色, 因为模的概念当中包含了交换群的概念. 虽然限制到实向量空间带来的简化远比这种限制必然导致的一般性的缺失有价值, 但是恰恰是**无需附加的努力就能够**学习到越来越普遍的结果的这种可能性, 使得年轻人快速地达到百年来数学研究的前沿, 而这段时间内人们的研究发现硕果累累.

没有**习题**的书对于初学者来说是要影响其使用的. 在本书中有几百个习题, 分成三类. 一些是课文中陈述的理论的实际的甚至是数值的举例, 初学者不解决相当数量的这类习题就不可能获得计算技术. 另外一些习题是课文的基本理论的补充, 演练这些习题, 读者将会习惯于运用课文中所使用的语言和推理模式; 这类习题中不是**非常**容易的, 就在前面标注 ¶. 最后一类是对于课文的重要的补充而且其解答是困难的, 这些习题专门针对那些真正对数学有兴趣的超前的学生; 这些习题前面冠以两个甚至三个 ¶.

我们认为, 解一个习题不能只靠匆匆打的一个 "草稿" 就自以为差不多明白了, 如果这样做对于解数值计算的习题是允许的, 那么对于偏重理论的习题, 必须**完整地**撰写出解答, 并且应当构造它们的真正的证明. 通过这种方式, 也只有通过这种方式, 学生们才能获得清晰而正确的语言, 按照专业术语的本意使用它们, 在数学里, 这是理解一个主题的最明确的标志.

最后, 与面向初学者的著作的传统相反, 我们向读者提供经过精心选择的书目组成的参考文献, 其中的大部分出自一流数学家. 读者得到并且使用一部分这样的书, 就能了解其他可能的观点和养成查阅参考书的习惯, 我们认为对于初学者这是大有裨益的.

1962 年 7 月

法兰西数学精品译丛

注：书号前缀为 978-7-04-0xxxxx-x

书号	书名	著者
★24308-6	解析函数论初步	H. 嘉当
★25156-2	微分学	H. 嘉当
★28417-1	广义函数论	L. 施瓦兹
★25801-1	微分几何	M. 贝尔热、B. 戈斯丢
★26362-6	拓扑学教程	G. 肖盖
★25155-5	谱理论讲义	J. 迪斯米埃
★24619-3	拟微分算子和 Nash-Moser 定理	S. 阿里纳克、P. 热拉尔
★29467-5	解析与概率数论导引	G. 特伦鲍姆
★33238-4	概率与位势（第Ⅰ卷）	C. 德拉歇利、P.-A. 梅耶
★31960-6	无穷小计算	J. 迪厄多内
★33238-4	广义系统的精确控制、摄动和稳定性（第一卷）精确控制论	J.-L. 利翁斯
★28757-8	代数学教程	R. 戈德门特
	概率与位势（第Ⅱ卷）	C. 德拉歇利、P.-A. 梅耶
	金融数学导引	El Karoui, E. Gobet
	完全集与三角级数	Jean-Pierre Kahane
	分析与代数原理 (及数论)	Pierre Colmez

说明：加★者已出版.

目录

第一章　集合论

§0 至 §5 的目的是引进集合与函数的概念, 没有这些概念, 我们在数学上什么也不能做. 反之, 使用这些概念, 我们能够做一切. 这些概念, 至少在本书所呈现的一般形式下, 在 19 世纪末之前还没有被剥离出来. 过去, 人们不明晰地谈论集合, 函数概念则涵盖了不同的对象; 代数函数, 解析函数, 可导函数, 连续函数, 等等, 单变量函数, 两个变量函数, 单复变量函数, 等等. 体现了数学的历史发展过程中所加的种种限制. 当今所有这些概念都是唯一的更加一般的集合概念的特殊情形, 这个概念观念上比它所包含的所有特殊情形更简单. 同时集合论的语言 (人们有时会修改其术语, 但不会修改基本概念) 正在广泛推广, 而是否应用集合论则变成了判断一个论述是否清晰和严格的条件.

下面几节对于阅读本书后续部分几乎是不可或缺的. §1 和 §2, §3 的第 1 小节立即就有重要的应用; 读者可以在需要 §4 时才认真研究它. 对于已经熟悉自然数的主要性质的读者来说, §5 则有点儿不太实用, 而本节的第 7 小节将经常被用到.

至于 §0, 这是数理逻辑的一个引论; 我们试图提供数学家构思他们所关心的对象的方式的大致想法, 并且在此汇集了一些特别重要的推理模式. 这一节, 跟 §1、§2、§3 一样, 开始不必仔细研读, 因为这里的概念经常被用到, 读者不但必然会逐渐熟悉它们, 而且在多数情况下还会很快地熟悉它们.

最后我们建议初学者不要惧怕这里的艰涩的外表, 尽管前几节无疑是极其抽象的. 给初学者的最好的劝告是完全忘记他可能已经了解的数学 (特别是整个初等几何, 除了 “几何变换” 的一般概念, 它和这里所处理的课题无任何关系). 还建议大家准确无误地援引专业术语的定义.

§0　逻辑推理

1. 逻辑完美的构思

在数学里有三个基本的过程: 构造数学对象, 建立这些对象之间的关系和证明这些关系中一些是真的, 或人们所说的, 是定理.

数学对象是数, 函数, 几何图形, 以及数学家所关心的无穷无尽的其他的东西: 这些对象不存在, 确切地说在自然界不存在, 但它们是程度不同地复杂且可见的物理学对象的抽象模型. 关系是可以用这些对象表达的 (真的或假的) 断言, 并且对应对于以数学对象作为其模型的自然对象所假设的性质. 至于真的关系, 对于数学家而言, 这是可以逻辑地一劳永逸地陈述的、从少量公理导出的关系, 这些公理用数学语言翻译人们所思考的具体对象的最 "显然的" 的性质. 而三段论的序列组成了给定定理的证明, 通过该序列从公理 (更实际地, 或已经建立的定理) 过渡到该定理.

这种解释, 对于一些初学的读者或许是贴切的, 而长期以来数学家已经不再满意, 这不仅仅是因为他们对于含糊不清的语句兴趣索然, 尤其是因为数学本身迫使他们必须小心地审视自己的科学基础, 用公式代替泛泛而谈, 而公式的意义容不得丝毫混淆, 并且可以用几乎是机械的方式就决定它是否是真的, 以及是否是有用的.

历史上, 把数学建立在尽可能牢固的基础上的必要性在 "集合论" 的发展过程中, 以及在数学中引进诸如环、域、群等 "抽象的" 新概念的过程中表现了出来.

涉及 Cantor 大约在 1870 年创立的集合论, 人们很快发现, 在这个理论里, 求助 "直观几何" 往好里说是无用的, 往坏里说其实是有害的. 二十年后人们发现要面对似乎与正常思维相悖但却被严密证明了的结果 (例如 Peano 证明了存在通过正方形所有点的连续曲线), 同时还要面对集合论内部实实在在的矛盾, 这些矛盾是由于不合时宜地使用数学家确信是谬误而又无法证明其逻辑错误的奇妙推理带来的. 而在数学里, 最可怕的莫过于内部矛盾, 因为希腊人已经知道, 如果拥有一个矛盾关系 (即同时为真的和假的), 那么可以直接证明所有其他的关系同样如此 (见后面的注 5).

至于所谓 "抽象的" 或 "公理的" 首批理论, 差不多处于同一时代 (1890—1910), 其目的一方面是把已经知道的一些特殊理论整合在一个一般理论里, 以便能够把研究这些理论以往用过的方法应用到另一些理论里, 另一方面是把从逻辑结构上看不够完善的理论建立在牢固的基础上 (后一种情形的最著名的例子是 Hilbert 所做的初等几何基础的研究. 希腊人之后两千年, 终于第一次有了几何的严密且纯粹演绎的陈述, 在这个叙述中, 所有公理无任何例外地被清晰地列出, 在那里人们清楚地看到, 对于每一个定理, 哪些公理对于它的成立是必需的, Hilbert 的陈述, 由于其语言和推理的严密性, 由于他拒绝所有的让步, 成了现代数学陈述的典范, 并且毫无疑问在许多世纪之内都将保持这个地位).

数学家们在紧密联系的这两个方面所做的努力, 使得他们习惯于强化他们在逻辑严密性方面的要求, 并且针对表面上距离具体“现实”越来越远的对象进行推理. 因此人们得以确信数学应予考虑的仅仅是汇集在一起的遵守若干清晰表述的“游戏规则”的符号, 它们组成数学的对象和关系.

今日人们甚至承认理论上可以仅仅使用少数几个*基本符号* (例如表示初等的逻辑运算的符号和两三个纯数学符号 —— 下节就要引进的相等和属于号, 或许还有 §2 将引进的构成对象的“序偶”的符号) 和无限量的*字母*书写出整个数学. 在数学的这种设想中, 数学对象和关系是由基本符号和字母按照某些一劳永逸地陈述出来的规则组成的汇集 (一般说来, 实际写出它们不成问题. 好奇的读者在本节的第 9 小节找到汇集的列表, 或更确切地说, 其中的一个列表). 在这个汇集中, 基本符号的作用是把某些逻辑的或数学的初等运算符号化, 同一个汇集内的符号的重复使用导致更复杂的运算; 而字母则被看作完全不确定的数学对象, 在所考虑的汇集里有引进“自由度”的效果, 也即形成与“任意变量”有关的关系和对象.

一旦建立了基本符号的列表和数学对象与关系的组成规则的列表, 剩下的任务就是陈述公理 (一些是纯逻辑的, 另一些则是严格数学性质的).

目前, 组成规则和公理已被选取和精确陈述, 以致证明定理和撰写数学文稿的机器的构思不再是空想. 描写这种机器 (*严密地并且专一地利用*一劳永逸地陈述的规则, 不忽略任何中间推理, 使它是“显然的”, 仅使用字母和基本符号, *不用任何通常的简化*) 的数学称为*形式化的*. 这种数学显然只在数学家的想象里存在. 一个用形式化语言写成的数学文本像很难懂的失败的报告, 却给能够看得懂它的读者以逻辑完美的感觉.

2. 数学的真实语言

人们算过, 如果试图用形式化语言写出即使像数 1 (表面看来) 这样简单的一个数学对象, 也得需要一个几万个符号 (基本符号的数目很小, 但同一个汇集里的每个符号自然可以重复多次) 的汇集. 试图处理类似汇集的数学家颇像登山家, 后者为了在峭壁上选择支撑点, 用电子显微镜考察峭壁.

在实践当中我们使用一群*缩写* (例如希腊字母 π, 符号 $+$, 普通语言的单词, 例如“数”, “点”, “直线”, “函数”, 等等), 这些缩写是为了用新的简单符号表示复杂的、由字母和基本符号组成汇集, 这些汇集还可能包含已经引进的缩写. 当人们充分地引进了基本符号的汇集的缩写, 继之缩写的汇集的缩写, 再继之缩写的汇集的缩写的汇集的缩写, 如此继续下去, 数学家不会设想(*) 如此构成的对象的完全和详尽的定义. 数学家现在关注的只是这一个复杂的等级是如何从*紧接着的*上一个复杂等级 (用通常的话说就是用所考虑的缩写构成的*定义*的那些内容) 过来的, 而不打算一步一步降到形式化语言的水平. 最终, 常常是这样的情况: 人们只对引进的缩写进

(*) 甚至应当说: 不再能想.

行工作, 而只要这些缩写的原始符号是与形式化语言的基本符号有相同的基础. (还是 Hilbert, 在他的几何基础的研究中, 首先引进三个本原概念 —— 点、直线和平面 —— 根本不定义它们, 而只建立应用它们的规则.)

自然在汇集的选取中, 数学家们的活动显著区别于机器的活动, 他们对于汇集的选择取决于用缩写所做的表示, 并且取决于对于什么感兴趣. 数学机器或许以光速推理, 但它可能仅局限于像布朗运动似地堆积定理. 相反, 数学家的目的则在于证明 "令人感兴趣的" 定理, 解决十几年来甚至几个世纪来提出的问题(*), 而非发现新的与已经提出的问题无关的毫无动机的数学分支. 事实上, 正是这些问题的研究迫使数学家提出新的概念 (即引进新的汇集) 和新技术, 而在初学者和使用者眼里, 这些有时看起来使得数学家远离了原来的问题. 例如, 在尝试证明 "Fermat 大定理" 和解决其他算术问题时, Kummer 大约在 19 世纪中期引进了 "理想数" 这一概念, 这个概念引导 Dedekind 约在 1870 年引进 "代数整环" 和这些环的 "理想" 的概念, 由此产生了环和理想的 "现代的" 和 "抽象的" 概念, 这些概念正处在和代数几何的融合之中. "Fermat 大定理" 说的是, 对于 $n \geqslant 3$, 方程为

$$x^n + y^n = 1$$

的 "平面曲线" 除它与坐标轴的交点外不包含平面的带有理坐标的点, 而这正是代数几何所关心的问题! 这个例子以及这里无法谈及的其他例子指出, 不管爱好者相信与否, 专业数学家致力于进入尽可能自然的道路, 在这条道路上, 他们的几何的或解析的或算术的直观 —— 因为存在算术的直观 —— 可以发挥作用.

数学家本人的活动还在另一个方面区别于机器的活动: 有生命的数学家最不像机器 (如果存在的话) 那样严格地符合绝对逻辑严密性. 在实践中, 最优秀的数学著作会包含一堆 "漏洞", 缺乏这些漏洞, 阅读这些著作将是难以忍受的. 这些逻辑上的漏洞并不重要, 因为每个人都相信如果愿意的话可以填补这些漏洞 —— 事实上, 初入门的读者甚至很可能觉察不到这些漏洞. 人们现在估计一个数学著作是 "完全" 正确的, 如果它达到了在初等数论的陈述中总有的清晰和严密的程度 (因此特别遗

(*) "化圆为方" 问题和 "借助直尺和圆规可以实现的" 几何作图的特征问题, 希腊人已经提出了, 仅到 19 世纪才得以解决. 在 18 世纪陈述的哥德巴赫猜想 (即所有偶整数可以写成两个奇素数的和) 至今还没有被证明, 尽管在这个道路上三十年来已经有了巨大的进步. 华林问题 (指出对于所有整数 n, 存在一个整数 p, 使得所有的整数都可以写成 p 项的和, 每项都是整数的 n 次幂 —— 最简单的情形是 $n=2$, 人们指出, 所有整数是四个平方数的和) 也是在 18 世纪提出, 而被 Hilbert 在 20 世纪初解决. "Fermat 大定理" (指出如果 n 是至少为 3 的整数, 则方程

$$x^n + y^n = z^n$$

没有任何由三个严格正的整数 x, y, z 组成的解) 在 17 世纪初宣布, 还远远没有解决 (这个问题已经在 1994 年 9 月由英国数学家怀尔斯解决 —— 译者注), 它无疑要求动用原来由代数学家在一个世纪里所发现的整个武器库和其他更有威力的技术. 我们这里必然只介绍几个问题, 他们的陈述充分简单, 以便初学者理解.

憾的是, 数学的这个分支在法国的中等教育中已经没有地位). 现在大量数学理论能够以这种风格陈述, 而这种可能性的一个后果是, 现今人们不再接受描述性的陈述, 而就在不久之前, 某些数学家还靠这类陈述同时在法国科学院和法兰西语言研究院谋取院士席位.

我们现在尝试给读者提供关于构造关系的方式和最重要的逻辑推理的更明确的信息. 我们希望初学者千万不要以为以下的论述仅使哲学家和数理逻辑专家感兴趣, 它事实上涉及的是推理的规则, 数学家每时每刻都在使用它, 而经验表明初学者在这方面是屡犯错误的.

3. 初等逻辑运算

正如我们在第 1 小节说过的, 关系和数学对象从理论上说就是按照某些规则组成的字母和基本符号的汇集, 读者将逐步接触这些基本符号. 首先感兴趣的是两个最简单的用以构成关系的符号和用这两个符号表示的一些缩写. 逻辑学家用 $\vee$ 和 $\neg$ 表示这两个符号, 数学中总会用到它们. 如果 R 和 S 是关系, 那么汇集

$$\mathrm{R} \vee \mathrm{S}$$

仍然是关系, 它由先写汇集 R, 再写符号 $\vee$, 最后写汇集 S 而得到, 称这个关系为关系 R 和 S 的**逻辑析取** [后面会看到, 当定义了关系 "真的" 以后, 关系 $(\mathrm{R} \vee \mathrm{S})$ 是真的, 只需给定的两个关系中至少一个是真的]. 同样, 如果 R 是一个关系, 则汇集

$$\neg\mathrm{R}$$

还是一个**关系**, 这个汇集由在汇集 R 前冠以符号 $\neg$ 而得到, 这个关系称为关系 R 的**否定** [后面将会看到, 说一个关系是假的, 如果它的否定是真的]. 这两个符号的应用规则 (即涉及它们的公理) 后面将予叙述, 暂时还不需要他们.

从这两个符号出发, 可以引进经常使用的缩写, 首先是符号

$$\wedge;$$

给定关系 R, S 之后, 用 $\mathrm{R} \wedge \mathrm{S}$ 表示关系

$$\neg[(\neg\mathrm{R}) \vee (\neg\mathrm{S})],$$

称为 R 和 S 的**逻辑合取** [后面将会看到, 关系 $(\mathrm{R} \wedge \mathrm{S})$ 是真的, 必须且只需给定的两个关系 R, S 都是真的]. 另外, 把关系

$$\mathrm{S} \vee (\neg\mathrm{R})$$

记成

$$\mathrm{R} \Rightarrow \mathrm{S}$$

称这个关系为**逻辑蕴含**, 并且读作[(*)]

$$\text{R 蕴含 S;}$$

[正如将要看到的一样, 这个关系是真的蕴含 S 是 R 的逻辑推论 (从而如果 R 是真的, 则 S 是真的), 但是关系 (R ⇒ S) 必然是真的, 如果 R 和 S 都不是真的. 例如, 关系

$$(1=2) \Rightarrow (2=3)$$

显然是真的, 同样 (在纳粹逻辑里) 关系

共产主义者焚烧了德国国会大厦, 故必须灭绝共产主义者

也如此. 纳粹逻辑的推断是无可厚非的, 但其前提是假的].

最后, 用

$$R \Leftrightarrow S$$

表示

$$[(R \Rightarrow S) \wedge (S \Rightarrow R)],$$

并且称为**逻辑等价**, 读作

$$\text{R 等价于 S.}$$

后面还有其他形成关系的规则.

4. 公理和定理

前一小节引入的符号已经让我们能够定义**真的**关系或**定理**, 它们可以通过重复利用下列两个规则而得到:

(RV1) *通过应用一个公理而得到的关系是真的.*

(RV2) *给定了关系 R 和 S 之后, 如果关系 (R ⇒ S) 是真的, 并且关系 R 是真的, 则关系 S 是真的.*

至于**公理**, 或者是一劳永逸地清晰陈述的关系, 或者是涉及 "任意的" 关系的规则, 应用这些规则从指定的关系引导出另外的指定的 (按照定义, 真的) 关系; 下一小节将陈述的公理是上面第二种类型的.

一个关系称为**假的**, 如果它的否定是真的.

[(*)] 在实践中, 仅当 "R ⇒ S" 在后面精确阐明的意义下为真时才说 "R 蕴含 S". 反之至少目前在本节, 我们写这种关系而不关心它们是真还是假.

此外, 我们注意许多人现在使用符号 ⇒ 作为 "蕴含" 这个词的缩写, 大部分专业数学家不同意这种非中文的形式 (如果他们是法国人, 就是非法文的形式) 的写法. 在实践中正确使用符号 ⇒ 十分困难.

注 1 我们因此看到, 真的关系的特征是它们能够被证明, 而不是, 例如, 被抽象化了的通过实验可以验证的自然规律. 这就是数学的真和实验科学的真的本质不同. 在日常实践中, 初学者的总是对存在的事实 (指实验科学证明是真的事实) 感兴趣, 而不管数学是否对它真感兴趣.

注 2 显而易见, 在数学里判断一个给定的关系是否是真的这个问题依赖起初采用的公理系统: 在一个公理系统里一个真的关系可能在看起来与前一个同样 "自然的" 另一个系统里不再是真的. 例如, 人们构造了集合论的公理系统, 在其中 "存在无穷集合" 这个断言不是真的 (即不能被证明 —— 这就说明为什么这个断言是被数学家应用的集合论基础的一个公理).



此外还要注意, 一个非真的 (即不能被证明的) 关系未必是假的: 如果 R 是一个关系, 理论上很可能出现这种情况, 起初采用的公理不足以证明 R 和 ¬R, 称这样的关系为 (在所考虑的公理系统内) *不可决定的*. 最近人们可以证明在建立在 §0、§1 和 §2 的公理基础上的通常的数学里, 存在这样的关系 (见 §5, 第 5 小节的注 9). 如果 R 是一个这样的关系, 如果愿意的话, 人们就有权在基础公理的列表里补充 R 或 ¬R.

注 3 除了真的关系, 假的关系, 不可决定的关系, 原则上还应当考虑*矛盾的*关系, 即同时是真的和假的关系. 所有的人都希望数学基础公理系统中的各个公理之间是相容的, 即禁止存在矛盾的关系 —— 但人们还不能证明这一点. 如果矛盾的关系出现, 就必须舍弃或减弱基础公理系统中的某些公理.

5. 逻辑公理和重言式

到目前为止我们还没有陈述任何公理, 在本小节我们要陈述它们中的最简单的公理, 用它们来验证最初等的逻辑推理 (三段论, 双重否定, 等等). 这些公理有四个, 但是如果读者以为这里采用的阐述方式是唯一可能的, 那就错了: 存在许多从少数几个公理出发推导出逻辑推理规则的其他可能的方式.

(AL1) *如果 R 是一个关系, 则关系*

$$(\mathrm{R} \vee \mathrm{R}) \Rightarrow \mathrm{R}$$

是真的.

因此如果 R ∨ R 是真的, 那么根据前小节的规则 (RV2), R 就是真的.

(AL2) *如果 R 和 S 是两个关系, 则关系*

$$\mathrm{R} \Rightarrow (\mathrm{R} \vee \mathrm{S})$$

是真的.

因此, 如果 R 是真的, 则 (R ∨ S) 必定是真的, 这符合 "或" 这个词的通常意义.

(AL3) *如果 R 和 S 是关系, 则关系*

$$(\mathrm{R} \vee \mathrm{S}) \Rightarrow (\mathrm{S} \vee \mathrm{R})$$

是真的.

组合前两个公理, 我们看到, 如果 R 是真的, 则 $(\mathrm{S} \vee \mathrm{R})$ 也是真的, 或同样的, 如果 S 是真的, 则 $(\mathrm{R} \vee \mathrm{S})$ 是真的. 由于我们已经建立如果 R 是真的, 则 $(\mathrm{R} \vee \mathrm{S})$ 必定是真的, 我们得到如果关系 R 和 S 至少有一个是真的, 则 $(\mathrm{R} \vee \mathrm{S})$ 是真的.

(AL4) *如果 R, S 和 T 是关系, 则关系*

$$(\mathrm{R} \Rightarrow \mathrm{S}) \Rightarrow ((\mathrm{R} \vee \mathrm{T}) \Rightarrow (\mathrm{S} \vee \mathrm{T}))$$

是真的.

于是如果 R 蕴含 S, 即关系 $(\mathrm{R} \Rightarrow \mathrm{S})$ 是真的, 则关系 $(\mathrm{R} \vee \mathrm{T})$ 蕴含 $(\mathrm{S} \vee \mathrm{T})$. 这类陈述对于初学者似乎是平凡的, 但它们受到重视是因为可以从中得到内容丰富的推论. 事实上, 应当把前面的四个公理看作符号 "$\vee$" 和 "$\wedge$" 的机械的使用规则, 而非深刻的数学发现.

仅仅使用前面的公理可以证明的定理称为**重言式**. 数学家在其逻辑推理过程中大部分时间使用它们, 但并不明确地申明引用了它们. 这里给出几个重言式, 我们不全部证明它们.

(TL1) *如果 R, S 和 T 是关系, $(\mathrm{R} \Rightarrow \mathrm{S})$ 和 $(\mathrm{S} \Rightarrow \mathrm{T})$ 是真的, 则 $(\mathrm{R} \Rightarrow \mathrm{T})$ 是真的.*

这里作为例子给出证明. 应用 (AL4), 并且在其中用 S, T 和 $(\neg \mathrm{R})$ 分别代换 R, S, 和 T, 我们发现关系

$$(\mathrm{S} \Rightarrow \mathrm{T}) \Rightarrow [\mathrm{S} \vee (\neg \mathrm{R}) \Rightarrow (\mathrm{T} \vee (\neg \mathrm{R}))]$$

是真的. 考虑到符号 $\Rightarrow$ 的意义, 这表明关系

$$(\mathrm{S} \Rightarrow \mathrm{T}) \Rightarrow [(\mathrm{R} \Rightarrow \mathrm{S}) \Rightarrow (\mathrm{R} \Rightarrow \mathrm{T})]$$

是真的. 根据假设, 关系 $\mathrm{S} \Rightarrow \mathrm{T}$ 是真的; 规则 (RV2) 指出关系

$$(\mathrm{R} \Rightarrow \mathrm{S}) \Rightarrow (\mathrm{R} \Rightarrow \mathrm{T})$$

是真的. 由于根据假设关系 $(\mathrm{R} \Rightarrow \mathrm{S})$ 是真的, 重新应用规则 (RV2) 即发现 $\mathrm{R} \Rightarrow \mathrm{T}$ 是真的, 这就完成了证明.

(TL2) *如果 R 是一个关系, 则关系 $(\mathrm{R} \Rightarrow \mathrm{R})$ 是真的.*

事实上, 根据 (AL1) 和 (AL2) 关系

$$\mathrm{R} \Rightarrow (\mathrm{R} \vee \mathrm{R}), \quad (\mathrm{R} \vee \mathrm{R}) \Rightarrow \mathrm{R}$$

是真的, 剩下的是应用 (TL1).

注 4 考虑到符号 $\Rightarrow$ 的定义, 上面的陈述意味着对于任意的关系 R, 关系



$$\mathrm{R} \vee (\neg \mathrm{R})$$

是真的. 这推不出 R, ¬R 中至少有一个是真的 —— 这正是是否存在不可决定的关系的问题! 事实上, 人们已经建立了一个关系 $(\mathrm{R} \vee \mathrm{S})$, 它是真的, 但是不能从它直接推出关系 R, S 中的一个是真的.

(TL3) *如果 R 是一个关系, 则关系*

$$\mathrm{R} \Leftrightarrow \neg(\neg \mathrm{R})$$

是真的.

说 R 是真的归结为说 ¬R 的否定是真的, 即 ¬R 是假的.

(TL4) *如果 R 和 S 是关系, 则关系*

$$(\mathrm{R} \Rightarrow \mathrm{S}) \Leftrightarrow ((\neg \mathrm{S}) \Rightarrow (\neg \mathrm{R}))$$

是真的.

为了证明 R 蕴含 S, 证明 S 的否定蕴含 R 的否定是充分的 (和必要的). 反之, 陈述

$$(\mathrm{R} \Rightarrow \mathrm{S}) \Leftrightarrow ((\neg \mathrm{R}) \Rightarrow (\neg \mathrm{S}))$$

是假的, 并且是众多逻辑错误的来源 (给定了 "所有人都是要死的" 这个陈述, 由此可以证明 "所有狗是不死的"). 然而在日常生活中人们利用这个陈述的情况屡见不鲜, 并且有时夹杂着心理层面的理由: "右派人士支持 '法国的代数学', 从而左派人士支持 '独立的代数学' ", 这必然经常导致错误.

(TL5) *如果 R 和 S 是关系, 则关系*

$$(\mathrm{R} \wedge \mathrm{S}) \Rightarrow \mathrm{R}, \quad (\mathrm{R} \wedge \mathrm{S}) \Rightarrow \mathrm{S}$$

是真的; 进而如果 R 和 S 是真的, 则 $(\mathrm{R} \wedge \mathrm{S})$ 是真的.

由此推出, $(\mathrm{R} \wedge \mathrm{S})$ 是真的, *必须且只需*所考虑的 R 和 S 是真的, 这符合词 "与" 的朴素意义.

注 5 容易证明, 我们前面提及的, 如果存在矛盾的关系, 则所有其他的关系都是矛盾的.



事实上, (AL2) 和 (AL3) 指出

$$(\neg \mathrm{R}) \Rightarrow (\mathrm{S} \vee (\neg \mathrm{R}))$$

是真的, 由于根据假设 $(\neg R)$ 是真的, 关系

$$S \vee (\neg R), \text{ 即 } (R \Rightarrow S)$$

是真的. 而根据假设 R 是真的, 由此推出 S 是真的, 同样推出 $\neg S$ 是真的.

这个注是**归谬推理**的基础. 这种推理在于, 为了证明关系 R 是真的, 把 $(\neg R)$ 暂时加入到数学公理当中, 并且证明这样建立的 "新" 数学是矛盾的. 根据注 5, 所有的关系在新的系统里都是真的, 特别关系 R 是真的. 因此, R 是 (通常的) 数学公理和关系 $(\neg R)$ 的推论, 容易看出, 这意味关系

$$(\neg R) \Rightarrow R$$

(在通常的数学里, 即我们现在回到的原来的数学) 是真的. 剩下要证明 R 本身是真的. 在公理 (AL4) 中, 用 $(\neg R)$, R 和 R 分别代替 R, S 和 T, 就得到关系

$$(\neg R \Rightarrow R) \Rightarrow [(\neg R \vee R) \Rightarrow (R \vee R)]$$

是真的. 由于根据 (AL3) 和注 4, $((\neg R) \Rightarrow R)$ 是真的, 因此 $(R \vee R)$ 是真的, 而 (AL1) 最终保证 R 是真的.

在实践中, 归谬推理这样使用: "假定关系 R 是假的", 这正好归结为添加 $(\neg R)$ 到数学公理中. 由此出发人们推理直到发现一个同时是真的和假的关系, 继而说 "这是荒谬的, 从而 R 是真的".

另一个经常使用的推理方法是**情况析取法**, 它建立在下列陈述的基础之上:

(TL6) *设 R, S 和 T 是关系, 如果三个关系*

$$R \vee S, \quad R \Rightarrow T, \quad S \Rightarrow T$$

是真的, 则 T 是真的.

由于 S 蕴含 T, 公理 (AL4) 指出 $(S \vee R)$ 蕴含 $(T \vee R)$. 由于 R 蕴含 T, 同样看到 $(R \vee T)$ 蕴含 $(T \vee T)$. 根据 (AL3), $(T \vee R)$ 蕴含 $(R \vee T)$, 于是有 $(S \vee R)$ 蕴含 $(T \vee T)$; 而 $(R \vee S)$ 是真的, 并且蕴含 $(S \vee R)$; 于是 $(T \vee T)$ 是真的, 借助 (AL1) 就完成了证明.

在实践中, 人们尤其喜欢在使用这个陈述时取 S 为 R 的否定; 为了证明 T 是真的, 只需指出 R 蕴含 T, 并且 $\neg R$ 也蕴含 T.

6. 关系中的代换

设 R 是一个关系, A 是一个数学对象, 而 x 是一个字母 (它是一个完全未确定的数学对象). 在组成关系 R 的字母和基本符号的汇集里, 处处用 A 代替 x, 构成关系的准则之一是这样得到的汇集还是一个关系, 用记号

$$(A|x)R$$

表示(*), 称为**在 R 中用 A 代换** x 得到的关系, 或**在 R 中给** x **以值** A 得到的关系; 如果关系 $(A|x)R$ 是真的, 则说 A **满足关系** R. 显而易见, 如果在汇集 R 里, x 实际上不出现, 那么关系 $(A|x)R$ 就是 R, 这时说 A 满足 R 就意味着 R 是真的.

为了指出字母 x 出现在关系 R 中, 经常把它写成形式

$$R\{x\}$$

(类似于在 §2 中表示函数的记号), 因此经常用

$$R\{A\}$$

代替 $(A|x)R$. 同样的, 如果 x 和 y 是出现在 R 里的两个字母, 为了明确这一事实, 我们用

$$R\{x, y\}$$

代替 R, 以此类推.

(TL7) *设 R 是一个关系, x 是一个字母, 而 A 是一个数学对象. 如果关系 R 是真的, 则在 R 里用 A 代换 x 得到的关系还是真的.*

换句话说, 如果把其中的 x 看作一个 "未确定的对象" 时, R 是真的, 那么当给 x 指定一个值 A 时, R 仍然是真的. 比如, 如果人们证明了关系

$$x = x,$$

其中 x 这个字母表示有关名词的数学意义, 则关系

$$A = A$$

对于任何数学对象 A 是真的.

这个结果无疑给初学者以玩一词多义游戏的印象, 其原因是前面给予字母以朴素意义, 它被认为表示 "完全未确定的" 或 "任意的" 对象; *但是这个解释至此既没有建立在精确的数学结果之上*, 也实际上没有被支配字母使用的、并且符合大众对于 "未确定的" 对象的理解的规则所验证. 规则 (TL7) *的目的正是验证字母作为 "未确定的对象" 的这个解释*. 相信人们通过 "常识" 的简单考虑可以验证 (TL7), 完全是误解: 用以证明的机器并不知道这个概念, 人们耗费大量时间才能让它们懂得什么是 "未确定的对象".



为了证明 (TL7), 首先对于应用一个公理直接得到的 R 验证它. 在这种情形, 关系 $(A|x)R$ 或者等于 R, 或者是直接应用跟 R 同样的公理而得到的. 例如 R 是关系

$$(S \vee S) \Rightarrow S,$$

(*) 仅在这一节用之.

其中 S 是一个给定的关系, 那么显然 $(\mathrm{A}|x)\mathrm{R}$ 就是关系

$$(\mathrm{S}' \vee \mathrm{S}') \Rightarrow \mathrm{S}',$$

其中 S′ 表示关系 $(\mathrm{A}|x)\mathrm{S}$. 假定清晰地写出了所有公理, 并且一旦实现了所有这些验证, 就直接过渡到了 “任意” 一个真关系的一般情形.

7. 量词

到目前为止, 我们已经指出了构成关系的三个基本方法 —— 逻辑析取, 否定和一个对象代换一个字母. 这些方法是纯逻辑性质的 (即不涉及在 §1 和 §2 将要引进的数学符号). 在实践中还需要形成关系的第四个纯逻辑方法, 这个方法表达断言: 给定一个关系 R 和一个字母 x 之后, 至少有一个数学对象 A 使得关系 $(\mathrm{A}|x)\mathrm{R}$ 是真的, 即它满足 R. 自然, 它的纯直觉意义是 “存在”. 以下讨论的目的是用一个新符号

$$\exists$$

代替 “存在”, 并且把这个符号的使用规则系统化, 使得它的表现符合日常语言中使用 “存在” 这个词时所要表示的普通意义. 符号 $\exists$ 称为**存在量词**.

于是给定一个关系 R 和一个字母 x 之后, 可以组成一个新的关系, 用

$$(\exists x)\mathrm{R} \text{ 或 } (\exists x)\mathrm{R}\{x\}$$

表示, 读作(*) 存在 x 满足 R.

例如, 为了要用这个语言书写方程 $x^4+1=0$ 至少有一个实根 (这是假的, 但无关紧要), 我们就用 $\mathbf{R}$ 表示实数集, 用下节将要引进的符号 $\in$, 组成关系

$$(\exists x)[(x \in \mathbf{R}) \wedge (x^4+1=0)].$$

除了符号 $\exists$, 我们还引进符号

$$\forall,$$

称为**全称量词**. 如果 R 是一个关系, 而 x 是一个字母, 用

$$(\forall x)\mathrm{R}$$

表示关系

$$\neg[(\exists x)(\neg \mathrm{R})],$$

读作

$$\text{对于所有 } x, \mathrm{R}$$

(*) 在日常实践中正确使用符号 $\exists$ 和 $\forall$ 十分困难, 因此宁肯写成 “存在” 和 “对于所有”.

或

对于任意 x 有 R.

说关系 $((\forall x)\mathrm{R})$ 是假的, 意味着断言 $((\exists x)(\neg\mathrm{R}))$ 是真的: 断言 “所有的人是会死的” 是断言 “存在不会死的人” 的否定. 但是显然说断言 “阿尔及尔的加斯巴 (Casbah) 的所有居民在 1957 年遭受酷刑” 是假的, 并不意味断言 “阿尔及尔的加斯巴 (Casbah) 没有居民在 1957 年遭受酷刑” 是真的(*).

¶注 6(**) 仅使用字母, 三个逻辑符号



$$\vee, \quad \neg, \quad \exists$$

和下两节中的用来形成 “相等” “属于” 和 “序偶” (甚至可以取消最后一个) 的数学符号, 理论上就可以书写整个数学.

在这个系统里, 一个断言 $(\exists x)\mathrm{R}$ 可以是真的, 却没有任何办法 “实际构造” 一个满足 R 的数学对象 —— 同样, 在日常生活中, 断言 “存在诚实的银行家” 并不算有实质意义的信息, 因为它本身并不展示出一个诚实的银行家.

在第 9 小节将会发现一个更复杂的系统, 没有上述的困扰, 但与前面引入的系统相比显得不够自然.

8. 量词使用规则

第一个规则是

(AL5) *设 R 是一个关系, x 是一个字母, 而 A 是一个数学对象, 则关系*

$$(\mathrm{A}|x)\mathrm{R} \Rightarrow (\exists x)\mathrm{R}$$

是真的.

因此如果 $(\mathrm{A}|x)\mathrm{R}$ 是真的, 换句话说, 如果 A 满足 R, 则 $(\exists x)\mathrm{R}$ 是真的, 这符合我们的期望. 在实践中几乎总是这样: 为了证明形如 $(\exists x)\mathrm{R}$ 的关系是真的, 人们展示出一个满足 R 的特定对象 A. 日常生活中也如此, 证明存在诚实的银行家的最好方式, 莫过于明确展示出一个来给大家.

(TL8) *设 R 是一个关系, x 是一个字母, 而 A 是一个数学对象, 则关系*

$$(\forall x)\mathrm{R} \Rightarrow (\mathrm{A}|x)\mathrm{R}$$

是真的.

(*) 想知道详情的读者可以查阅 Pierre Vidal-Naquet 的书中的类似的文献, (该书书名《国家理智》, 由半夜出版社出版, 巴黎, 1962), 其中有丰富的参考文献. 关于阿尔及利亚警察的愚蠢行为, 参见同一作者的《层出不穷的问题》, 刊于 1965 年 9 月 29 日的《世界报》.

(**) 注 6 前面的 “¶” 表示这个注稍难, 可量力阅读; 如果注前加 “¶¶”, 表示这个注比较难, 初读可以不看. —— 译者注

于是如果关系 $(\forall x)\mathrm{R}$ 是真的, 则在 R 里用任何一个数学对象代换 x 得到的关系都是真的. 例如, 如果证明了关系

$$(\forall x)(x = x)$$

是真的, 其中 x 这个字母表示有关名词的数学意义, 则对于所有数学对象 A 关系

$$\mathrm{A} = \mathrm{A}$$

都是真的. 日常生活中人们如法行事: 断言 "所有脑力劳动者是同性恋者" 蕴含每一个明确点名的脑力劳动者是同性恋者.

(TL9) *设 R 是一个关系, x 是一个字母, 则关系*

$$\neg((\exists x)\mathrm{R}) \Leftrightarrow (\forall|x)(\neg\mathrm{R})$$

是真的.

说关系 $(\exists x)\mathrm{R}$ 是假的意味着关系 $(\forall x)(\neg\mathrm{R})$ 是真的, 特别地说, 所有数学对象满足 $(\neg\mathrm{R})$.

(TL10) *设 R 是一个关系, x 是一个字母, 则关系*

$$(\forall x)(\mathrm{R} \wedge \mathrm{S}) \Leftrightarrow ((\forall x)\mathrm{R} \wedge (\forall x)\mathrm{S})$$

是真的.

直觉上这是显然的. 反之注意关系

$$(\forall x)(\mathrm{R} \vee \mathrm{S}) \Leftrightarrow ((\forall x)\mathrm{R} \vee (\forall x)\mathrm{S})$$

一般不是真的. 比如, 断言 "所有脑力劳动者是变节者或同性恋者" 不蕴含 "所有脑力劳动者是变节者或所有脑力劳动者是同性恋者", 因为可能存在脑力劳动者, 他是变节者而非同性恋者, 同时存在脑力劳动者, 他既是同性恋者又是爱国者.

(TL11) *设 R 和 S 是关系, 而 x 是一个字母, 则关系*

$$(\exists x)(\mathrm{R} \vee \mathrm{S}) \Leftrightarrow ((\exists x)\mathrm{R} \vee (\exists x)\mathrm{S})$$

是真的.

直觉上这也是显然的. 注意关系

$$(\exists x)(\mathrm{R} \wedge \mathrm{S}) \Rightarrow ((\exists x)\mathrm{R} \wedge (\exists x)\mathrm{S})$$

是真的, 但是关系

$$((\exists x)\mathrm{R} \wedge (\exists x)\mathrm{S}) \Rightarrow (\exists x)(\mathrm{R} \wedge \mathrm{S})$$



一般是假的. 断言“存在富有并且诚实的人”显然蕴含“存在富有的人并且存在诚实的人”, 但是反向的蕴含是不正确的, 因为纯逻辑推理不足以排除下列可能性, 富有的人必然是不诚实的.

作为结束, 我们陈述几个有关重复量词的规则. 设 R 是一个关系, 而 x 和 y 是两个不同的字母; 我们可以在 R 里应用一个关于字母 x 的量词, 而后在所得到的关系里, 应用关于字母 y 的量词; 不言而喻人们可以按相反的次序进行, 并且问道执行运算的次序是否有其重要性. 回答由下面的陈述给出 (这不覆盖所有可能的情形, 理由很简单, 所有人想得到而下面没有出现的陈述都是错误的):

(TL12) *设 R 是一个关系, 而 x 和 y 是两个不同的字母. 则*

$$(\forall x)(\forall y)\mathrm{R} \Leftrightarrow (\forall y)(\forall x)\mathrm{R},$$
$$(\exists x)(\exists y)\mathrm{R} \Leftrightarrow (\exists y)(\exists x)\mathrm{R},$$
$$(\exists x)(\forall y)\mathrm{R} \Rightarrow (\forall y)(\exists x)\mathrm{R}$$

是真的.

比如给出第二个陈述的一个直觉 (即不正确的) 证明. 如果关系 $(\exists x)(\exists y)\mathrm{R}$ 是真的, 这就是说可以找到一个对象 A 使得在关系 $(\exists y)\mathrm{R}$ 里用它代换 x 会得到一个真的关系, 换句话说关系 $(\exists y)\mathrm{R}\{\mathrm{A},\mathrm{y}\}$ 是真的. 而这正说明存在一个对象 B 使得用它代换 $\mathrm{R}\{\mathrm{A},\mathrm{y}\}$ 中的 y 就会得到一个真的关系, 换句话说关系 $\mathrm{R}(\mathrm{A},\mathrm{B})$ 是真的. 而这时 $(\exists x)\mathrm{R}\{\mathrm{x},\mathrm{B}\}$ 是真的, 从而关系 $(\exists y)(\exists x)\mathrm{R}\{x,y\}$ 是真的, 而这就 (按直觉的方式) 证明了所考虑的关系.

注 7 这个证明至少有三处不正确: (1) 只证明了关系 $(\exists x)(\exists y)\mathrm{R}$ 蕴含 $(\exists y)(\exists x)\mathrm{R}$, 而原来要求的是证明这两个关系是等价的, 这个显然并不严重, 反向蕴含按相同的方式进行; (2) 我们只证明了如果关系 $(\exists x)(\exists y)\mathrm{R}$ 是真的, 则关系 $(\exists y)(\exists x)\mathrm{R}$ 也是真的; 而一个蕴含可以完完全全是真的, 却不依赖其两项的真实性; (3) 我们认为形如 $(\exists x)\mathrm{R}$ 这样的关系是真的, *必须且只需*可以找到一个对象 A 使得关系 $(\mathrm{A}|x)\mathrm{R}$ 是真的: 但是规则 (AL5) 仅指出这个条件是*充分的*.

关于 (3), 可以借助于第 9 小节的考虑来验证. 对于 (2), 我们使用**辅助假设法**来验证, 这个方法是: 设 R 和 S 是两个关系, 临时把 R 添加到公理系统里 (说“假定 R 是真的”以特别指出这一点), 假定 S 在“新”数学里是真的; 那么关系 $(\mathrm{R} \Rightarrow \mathrm{S})$ 是真的 (显然是在通常的数学里!). 这个方法借助第 4 小节和第 5 小节的更初等的准则来验证: 从“新”的公理 (即通常的公理和 R) 出发写出三段论推理序列, 最后到达 S, 通过一步步的推理证明链条中的每一个关系都是 R 的逻辑推论, 在过程的最后就是蕴含 $(\mathrm{R} \Rightarrow \mathrm{S})$.

注 8 经常会碰到写出依次出现的两个 (甚至更多) 量词的否定这类习题, 初学者做这类习题往往会难以完成. 例如, 写出关系 $(\forall x)(\exists y)\mathrm{R}$ 的否定 (或等价于否定的关系). 暂时用 S 表示关系 $(\exists y)\mathrm{R}$, 我们应当组成 $(\forall x)\mathrm{S}$ 的否定; 而

$(\forall x)$S 正是

$$\neg[(\exists x)(\neg \text{S})]$$

它的否定根据 (TL3) 等价于 $(\exists x)(\neg \text{S})$. 但是 S 表示的正是 $(\exists y)$R, 此关系等价于 $(\exists y)(\neg(\neg \text{R}))$, 我们看到 $(\neg \text{S})$ 等价于 $(\forall y)(\neg \text{R})$. 因此我们得到要找的结果, 即关系

$$(\forall x)(\exists y)\text{R}$$

的否定等价于关系

$$(\exists x)(\forall y)(\neg \text{R}).$$

9. Hilbert 运算, 组成准则

为了结束本节, 以及满足对于逻辑感兴趣的读者的意愿, 我们简略地给出一个可能的逻辑系统, 这个系统被 N. Bourbaki 在其数学基础中所应用.

在这个系统里, 使用七个基本符号和字母. 四个基本符号, 即

$$\vee, \quad \neg, \quad \tau, \quad \square$$

是纯逻辑符号性质的; 三个确切地说是数学性质的, 它们是

$$=, \quad \in, \quad \supset,$$

(第三个是序偶符号, 见下面的准则 (OM2)), 我们将在 §1 和 §2 引进它们.

写出符号和字母的一个系列得到一个汇集, 出现在汇集里的某些符号 τ 可以用直线连接到某些符号 $\square$ —— 例如, 表达式

$$\tau x \in y = \in \in yx = z\square$$

是一个汇集.

设 A 是一个汇集, 而 x 是一个字母. 我们要指出一个产生新汇集的过程, 新的汇集不再含有字母 x, 但仍然采用记号

$$\tau_x(\text{A})$$

表示它; 通过以下三个操作得到它:

a) 在 A 前写符号 τ 得到汇集 τA;

b) 用连接符连接 A 前面的 τ 和每个字母 x;

c) 在已经得到的汇集里, 处处用符号 $\square$ 代替 x.

比如, A 是前面写出的汇集, 那么 $\tau_x(\text{A})$ 是汇集

$$\tau\tau\square \in y = \in \in y\square = z\square.$$

从 A 过渡到 $\tau_x(\mathrm{A})$ 所做的运算本质上属于 Hilbert, 后面将给出它的直观意义.

现在叙述数学对象和关系的组成准则:

(OM1) 所有字母是一个数学对象.

(OM2) 如果 A 和 B 是数学对象, 则汇集

$$\supset \mathrm{AB}$$

是一个数学对象, 在实际中用

$$(\mathrm{A}, \mathrm{B})$$

表示它.

(OM3) 设 A 和 T 是数学对象, x 是一个字母, 则在 T 里处处用 A 代换字母 x 导出的汇集 $(\mathrm{A}|x)\mathrm{T}$ 是一个数学对象.

(OM4) 设 R 是一个关系, 而 x 是一个字母, 则汇集 $\tau_x(\mathrm{A})$ 是一个数学对象.

(R1) 设 R 和 S 是关系, 则汇集

$$\vee \mathrm{RS}$$

是关系, 在实际中把 $\vee\mathrm{RS}$ 写成 $\mathrm{R} \vee \mathrm{S}$.

(R2) 设 R 是一个关系, 则汇集

$$\neg \mathrm{R}$$

是一个关系.

(R3) 设 R 是一个关系, x 是一个字母, 而 A 是一个数学对象, 则汇集 $(\mathrm{A}|x)\mathrm{R}$ 是一个关系.

(R4) 设 A 和 B 是数学对象, 则汇集

$$= \mathrm{AB}$$

是一个关系, 在实际中把 $=\mathrm{AB}$ 写成 $\mathrm{A} = \mathrm{B}$.

(R5) 设 A 和 B 是数学对象, 则汇集

$$\in \mathrm{AB}$$

是一个关系, 在实际中把 $\in \mathrm{AB}$ 写成 $\mathrm{A} \in \mathrm{B}$.

没有其他的方法构成数学对象和关系. (OM4) 几乎从不直接使用, 除此之外, 前面的所有准则在实际中每时每刻都会用到.

人们注意到前面没有提到量词 $\exists$ 和 $\forall$: 这是因为 (使用 Hilbert 运算) 能够作为简单的缩写引入它们.

精确地说, 设 R 是一个关系, 而 x 是一个字母, 则按照定义

$$(\exists x)\mathrm{R}$$

是关系, 而关系

$$(\tau_x(\mathrm{R})|x)\mathrm{R},$$

通过在 R 里处处用 $\tau_x(\mathrm{R})$ 代换字母 x 得到. 因此 $(\exists x)\mathrm{R}$ 是真的, 必须并且只需对象 $\tau_x(\mathrm{R})$ 满足关系 R, 而这就给出了 Hilbert 运算的直观解释: 这个运算在于对于每个关系 R 和每个字母 x 一劳永逸地选择满足关系 $\mathrm{R}\{x\}$ 的一个对象 (如果它 "存在", 在相反的情形, $\tau_x(\mathrm{R})$ 是一个人们无话可说的对象). 不言而喻这个 "存在" 纯粹是虚构的: Hilbert 运算的益处在于给了一个完全不自然但却纯粹机械的实际构造一个对象的过程, 人们仅仅知道该对象满足预先规定的条件 (在该对象存在的条件下)(*). 现在人们还用它来代替选择公理 (§2, 注 7).

在日常的实践中, 利用 Hilbert 运算是完全例外的 (参见 §5, 注 1, 基数的定义), 它显然不能够推演出任何 "明晰的" 结果. 像是哲学的上帝, Hilbert 运算是难以懂得和不见踪影的, 但是它统治一切, 并且它的体现无处不在.

§0 习题

1. 设 R 和 S 是两个关系. 证明, 如果 R 是假的, 则关系 $(\mathrm{R} \Rightarrow \mathrm{S})$ 是真的. 能由此推出 S 是真的吗?

2. 设 R 和 S 是两个关系. 证明关系

$$[\mathrm{R} \wedge \neg\mathrm{R}] \Rightarrow \mathrm{S}$$

是真的.

3. 借助一个例子证明关系

$$(\forall x)(\exists y)\mathrm{R} \Rightarrow (\exists y)(\forall x)\mathrm{R}$$

一般是假的.

4. 设 R 和 S 是两个等价的关系, 而 T 是任意一个关系. 证明以下每个关系都是真的:

$$(\neg\mathrm{R}) \Leftrightarrow (\neg\mathrm{S}),$$
$$(\mathrm{R} \Rightarrow \mathrm{T}) \Leftrightarrow (\mathrm{S} \Rightarrow \mathrm{T}),$$
$$(\mathrm{T} \Rightarrow \mathrm{R}) \Leftrightarrow (\mathrm{T} \Rightarrow \mathrm{S}),$$
$$(\mathrm{R} \wedge \mathrm{T}) \Leftrightarrow (\mathrm{S} \wedge \mathrm{T}),$$
$$(\mathrm{R} \vee \mathrm{T}) \Leftrightarrow (\mathrm{S} \vee \mathrm{T}).$$

¶¶5. 证明下列关系, 其中 R, S 和 T 表示任意关系:

$$\mathrm{R} \Rightarrow (\mathrm{S} \Rightarrow \mathrm{R}),$$
$$(\mathrm{R} \Rightarrow \mathrm{S}) \Rightarrow [(\mathrm{S} \Rightarrow \mathrm{T}) \Rightarrow (\mathrm{R} \Rightarrow \mathrm{T})],$$
$$\mathrm{R} \Rightarrow [\neg\mathrm{R} \Rightarrow \mathrm{S}],$$

(*) 如果能够构造一个满足 R 的对象 A, 则 $\tau_x(\mathrm{R})$ 满足 R 这一事实, 只不过是公理 (AL5) 的一个重新表达.

$$(R \vee S) \Leftrightarrow [(R \Rightarrow S) \Rightarrow S],$$

$$(R \Leftrightarrow S) \Leftrightarrow \{(R \wedge S) \vee [(\neg R) \wedge (\neg S)]\},$$

$$(R \Leftrightarrow S) \Leftrightarrow \neg[\neg R \Leftrightarrow S],$$

$$\{R \Rightarrow [S \vee (\neg T)]\} \Leftrightarrow [(T \wedge R) \Rightarrow S],$$

$$[R \Rightarrow (S \vee T)] \Leftrightarrow [S \vee (R \Rightarrow T)],$$

$$(R \Rightarrow S) \Rightarrow \{(R \Rightarrow T) \Rightarrow [R \Rightarrow (S \wedge T)]\},$$

$$(R \Rightarrow T) \Rightarrow \{(S \Rightarrow T) \Rightarrow [(R \vee S) \Rightarrow T]\},$$

$$(R \Rightarrow S) \Rightarrow [(R \wedge T) \Rightarrow (S \wedge T)],$$

$$(R \Rightarrow S) \Rightarrow [(R \vee T) \Rightarrow (S \vee T)].$$

¶6. 设 R 和 S 是两个等价的关系, 而 x 是一个没有出现在 R 中的字母. 证明关系

$$(\forall x)(R \vee S) \Leftrightarrow (R \vee (\forall x)S),$$

$$(\exists x)(R \wedge S) \Leftrightarrow (R \wedge (\exists x)S)$$

是真的.

7. 设 R 和 S 是两个等价的关系, 而 x 是一个字母. 证明关系

$$[(\forall x)(R \vee S)] \Rightarrow [(\forall x)R \vee (\exists x)S],$$

$$[(\exists x)R \vee (\exists x)S] \Rightarrow [(\exists x)(R \vee S)].$$

8. 一伙嗜血者准备吃掉一个传教士. 为了最后一次向他显示他们尊重人类的尊严和自由, 嗜血者向传教士提出由他本人决定自己的命运, 为此只需做一个简短的宣示: 如果这个宣示是真的, 他将被火烤; 如果是假的, 他将被水煮. 传教士为了挽救自己的生命该说什么? (出自 Cervantes.)

9. 上校 X 指控教授 Y 有谋杀罪. 两周以后, 教授企图煽动谋杀上校. 上校有理由吗?

10. 叙述下列断言的否定的等价断言(*):

a) 所有的直角三角形有一个直角.

b) 所有监狱的所有犯人憎恨所有看守.

c) 对于所有非负整数 x 存在一个非负整数 y, 使得对于所有非负整数 z 关系 $z < y$ 蕴含关系 $z < x + 1$.

11. 考察下列断言之间的逻辑关系:

A: 所有的人是必死的.

B: 所有的人是不死的.

C: 没有一个人是必死的.

D: 没有一个人是不死的.

E: 存在不死的人.

F: 存在必死的人.

(*) 写出一个关系的否定的最简单的方式 (自然) 是在其前面冠以符号 "¬". 在本习题里这显然不是对读者所要求的.

¶12. 借助 Hilbert 运算证明, 如果 R 是一个关系, 而且一个字母 x 出现在 R 内, 那么字母 x 不再出现在关系内 $(\forall x)$R 和 $(\exists x)$R 内, 虽然表示这两个关系的记号内出现了 x.

这个非常简单的结果表明, 在记号 $(\forall x)$R 内, 字母 x 仅为了指出对于关系 R 进行的一个运算, 这个运算的结果中的一个就是消去关系 R 中的 x. 一个类似的现象在传统的记号

$$\int_0^1 f(x)dx$$

中出现, 其中的记号 x 不起任何作用, 在最后的结果中不出现. 这个习题解释了为什么在关系 $(\forall x)$R 中, 如果愿意可以用任意其他在 R 中未曾出现过的字母代替 x; 例如, 关系

$$(\forall x)(x \times y = x - z) \quad 和 \quad (\forall t)(t \times y = t - z)$$

不仅是等价的, 其实是相同的. 反之关系

$$(\forall x)(x \times y = x - z) \quad 和 \quad (\forall y)(y \times y = y - z)$$

是不同的.

¶13. 在火星上粗略地讲有两种政治观点: 右的和左的. 另一方面, 火星上的大学生分属两个社团: 火星大学生协会 (UPEM) 和火星大学生联盟 (FPEM). 已知左派大学生加入 UPEM, 证明 FPEM 是不过问政治的.

14. 考虑四个非负整数 m, n, p, q, 对于它们做如下假设:

a) 非负整数 m, p 和 q 是互素的;

b) m 除以 pq 的余数是 12;

c) $2n + 3$ 除以 n 的余数是 4;

d) 不存在满足关系 $x^4 + y^5 = p^3 - q^2 + m^6$ 的一对非负整数 x, y.

证明 n 是偶数.

15. 考虑下列两个断言:

a) "在同阿尔及利亚的居住部长 Robert Lacoste 先生的全面协议中, 我们赋予第十空降师确保阿尔及利亚和平和安全的责任. 这个师在三个月内赢得了阿尔及尔的战争, 而不用重机枪射击建筑物, 也不会有一架法国飞机向 Casbah 发射子弹." (摘自 Salan 将军的诉讼中的声明).

b) "结果是 '阿尔及尔战争' 对于参战的空降部队和对于法国来说, 成为一种代价惨重的胜利: 据估计, Casbah 的 80000 人口中有 30% 至 40% 的成年男子在 '战争' 的一个或另一个阶段被逮捕和审讯, 审讯采用的主要手段是省时而快速有效的拷打." (Edward Behr,《阿尔及尔时报》通讯员, "阿尔及利亚问题", W. W. Norton, New York, 1961). 这两个断言逻辑上是不相容的吗? (不要求决定它们是真的或假的.)

16. 利用规则 $\neg(\neg A) \Leftrightarrow A$ 简化下列句子 (摘自足球比赛的一篇报道):

"··· 没有发现任何运动员否定对方不被不尊重 ···"

对于下文回答同样的问题:

"我还要给你寄一个关于考试的评语. 我坚持提醒几个基本的原则. 我讽喻某些 '裁决' 或某些行为, 幸亏这些还不太多. 但是, 即使它们保留得还不多, 即使它们应当被合法地确认, 不乏否定所有工作质量的价值的人, 而学院的大多数教师都努力把工作做好."

17. 对于 "Pelléas et Mélisande" 的连续演出, 一个记者在下列两个草稿之间踌躇不决:

a) 角色 Mélisande 从来没有被饰演得如此好.

b) 如此年轻又披有如此美发的女歌唱家从来没有如此好地饰演 Mélisande. 这些称颂的话哪一个更强烈? (阐明 ¬A 和 ¬B.)

§1 相等和属于关系

1. 相等关系

前一节引入的符号是纯逻辑性质的, 用它们 "构成" 不仅针对数学的推理模式. 而在这一节引入的两个 "基本符号" (相等符号和属于符号), 则用以构造特别针对数学的关系和对象.

相等符号记作

$$=,$$

以下列方式用它构成关系: 如果 a 和 b 是数学对象 (或**集合**, 两个术语是同义的), 我们得到一个关系, 写成

$$a = b,$$

即依次写出用 a 表示的符号和字母的汇集, 等号 $=$, 以及用 b 表示的汇集. 直观地说, 这个关系意味着, 当它是真的时, 用 a 和 b 所表示的具体对象是 "相等的" —— 我们并不深究这个概念的哲学意义. 对于数学家而言, 问题不在于理解或让人理解符号 $=$ 的 "深刻意义", 而是知道如何运用它. 为了做到这一点, 只需参考下面的陈述, 它概括了书写等式时遵循的 "游戏规则"(*):

定理 1 我们有下列性质:

a) 对于所有 x, 关系 $x = x$ 是真的.

b) 对于任意 x 和 y, 关系 $x = y$ 和 $y = x$ 是等价的.

c) 对于任意 x, y, z, 关系 $x = y$ 和 $y = z$ 蕴含 $x = z$.

d) 设 u 和 v 是使得 $u = v$ 的对象, 而 $\mathrm{R}\{x\}$ 是一个含有字母 x 的一个关系; 则在 $\mathrm{R}\{x\}$ 中处处分别用 u 和 v 代换字母 x 得到的关系 $\mathrm{R}\{u\}$ 和 $\mathrm{R}\{v\}$ 是等价的.



注 1 定理的断言 d) 意味着两个相等的对象 a 和 b 有同样的性质, 即所有对于 a 有效的断言对于 b 也是有效的, 反之亦真. 至于断言 a), b) 和 c), 它们表达的是等式的最 "显然的" 性质.

至于前述定理的证明, 仅用一张纸, 就能证明涉及等号 "=" 的定理, 简直就是奇迹. 事实上, 断言 d) 是数学基础公理中的一个公理, 至于 "显然的" 断言 a),

(*) 从现在起, 我们不再使用 §0 中的相对 "形式化" 的语言. 如果想使用它, 必须以下列方式陈述定理 1 的断言 a): 设 x 是一个字母, 则关系 $(\forall x)(x = x)$ 是真的. 定理的其他断言也写成类似的形式. 需要提醒的是, 在数学里人们不使用形式化语言 (其实这将是不可能的), 但我们还必须调整得不要背离这种写法太远, 以致如果有合适的理由需要时, 人再也回不去了.

b) 和 c), 或者把它们取作公理 (初学者应当这样做), 或者从一个更复杂的 (但是唯一的, 初学者不必深究) 公理推出, 这个公理是: 如果 R 和 S 是两个等价的关系, 而 x 是一个字母, 则关系

$$\tau_x(\mathrm{R}) = \tau_x(\mathrm{S})$$

是真的.

对于渴望深究数学基础的读者来说, 证明定理 1, 并且确信他发现的所有证明都是不充分的, 这将是一个绝好的习题.

当然, 在实际中, 我们常常使用定理 1, 而从不明确声明引用了它.

2. 属于关系

数学的第二个基本符号是**属于号**, 记作

$$\in,$$

像符号 "=" 一样, 用来从数学对象出发构造关系: 如果 a 和 b 是数学对象, 写出

$$a \in b$$

就得到一个关系, 读作

$$a \text{ 属于 } b$$

或

$$a \text{ 是 } b \text{ 的一个元素}.$$

关系 $a \in b$ 的否定记作

$$a \notin b.$$

这里同样, 唯一重要的事情是支配符号 $\in$ 使用的公理, 实际上只有一个, 这就是

定理 2 设 A 和 B 是两个集合, 为了有 A = B, 必须并且只需关系

$$x \in \mathrm{A} \quad \text{和} \quad x \in \mathrm{B}$$

是等价的.



注 2 容易给出符号 $\in$ 的直观解释, 使得定理 2 变得 "显然". 为此, 必须想象每个数学对象是另外一些对象的集体 (集合一词由此而来), 关系 $x \in y$ (当它是真的时) 意味着 x 是组成集合 y 的对象中的一个. 而定理 2 肯定, 为了对象的两个集体是相等的, 必须且只需它们含有同样的对象 (即第一个集体的所有对象属于第二个, 反之亦然).

在实践中, 人们经常想象数学对象或者是另一些对象的集体, 像我们刚才所说的, 或者是 “个体的” 的对象 (在第 6 小节将看到在这两个解释之间并无矛盾). 这两种情形的 “自然” 解释依赖于背景, 而读者少许动一下脑子, 就能容易地看出来. 一般地说, 当我们设想一个数学对象作为一个 “集合” 时, 就用大写字母表示它, 反之当把它看作一个集合的 “元素” 时, 就常用小写字母表示它 —— 在定理 2 的陈述中就是这样做的. 这个规则仅仅是一个习惯, 并且容许众多例外.

注 3 人们可以具体解释数学对象为集体或者另外一些对象的集合, 这个事实显然与数学地 (毕竟证明机器本身没有物理直观) 定义集合概念的问题毫不相干. 这个问题只有通过 §0, 第 9 小节的方法或类似的方法解决.



在中学生使用的一些课本中发现下列表述: “人们称所有同样性质的对象的集体为集合”. 对于这个 “定义” 的第一个异议在于它把词 “集合” 翻译成 “集体”; 而这两个词显然是同义词, 人们面对一个简单的同义词游戏. 第二个异议是这些课本的作者不认为合并任意两个 “集体”, 例如, 土豆的集体和梨的集体, 组成一个 “集体” 会有任何困难, 难道土豆和梨是同样性质的!

这个例子明显指出, 为了教学法的理由, 试图给集合 (或集体) 这个词以初等定义会走向荒诞. 人们宁肯把集合概念看作不定义的 (所有人直觉地理解) 原始概念, 借助它可以构成各种关系, 人们可以对它们进行逻辑推理.

3. 一个集合的子集

从属于符号出发, 我们要引进一个缩写, 记作

$$\subset$$

称为**包含符号**. 给定两个集合 A 和 B, 用

$$\mathrm{A} \subset \mathrm{B}$$

表示下列关系:

$$\text{对于所有 } x, \text{关系 } x \in \mathrm{A} \text{ 蕴含关系 } x \in \mathrm{B}.$$

换句话说, $\mathrm{A} \subset \mathrm{B}$ 意味着 A 的所有元素也在 B 里.

关系 $\mathrm{A} \subset \mathrm{B}$ 读作 A **包含于** B, 或 B **包含** A, 或 A **是** B **的一个子集**, 这个关系还可以写成

$$\mathrm{B} \supset \mathrm{A}.$$

定理 3 *包含关系具有下列性质:*

a) *关系* $\mathrm{A} \subset \mathrm{B}$ *和* $\mathrm{B} \subset \mathrm{C}$ *蕴含* $\mathrm{A} \subset \mathrm{C}$;

b) *为了有* $\mathrm{A} = \mathrm{B}$, *必须且只需* $\mathrm{A} \subset \mathrm{B}$ *和* $\mathrm{B} \subset \mathrm{A}$.

断言 a) 意即: 如果关系 $x \in \mathrm{A}$ 蕴含 $x \in \mathrm{B}$, 而且 $x \in \mathrm{B}$ 蕴含 $x \in \mathrm{C}$, 则这三个关系中的第一个蕴含最后一个 —— 逻辑上, 这正是 "三段论原则" 或 §0 的 (TL1). 断言 b) 显然归结为定理 2.

定理 4 设 $\mathrm{R}\{x\}$ 是出现字母 x 的一个关系. 对于所有集合 X, 存在 X 的唯一子集 A 具有下列性质: 为了有 $x \in \mathrm{A}$, 必须且只需关系 $x \in \mathrm{X}$ 和 $\mathrm{R}\{x\}$ 是真的.

我们说 A 是满足关系 $\mathrm{R}\{x\}$ 的 $x \in \mathrm{X}$ 的集合.

例 1 取 X 为自然数集合, 而 $\mathrm{R}\{x\}$ 是关系 "x 可以被 2 整除", 那么 A 就是偶整数的集合.

注 4 直观地, A 是由具有关系 $\mathrm{R}\{x\}$ 所表示的性质的对象 $x \in \mathrm{X}$ 组成的集体, 从而 A 的存在性直观地看来是显然的. 数学上说来, 不使用远不如它显然的公理, 我们不能证明定理 4. 因此建议初学者可以假定它是公理.



¶注 5 虽然有悖常理, 对于所有关系 $\mathrm{R}\{x\}$, 存在一个其元素是所有使得关系 $\mathrm{R}\{x\}$ 为真的对象 x 的集合 (在 §0 的精确意义下) 这一断言不是真的 (定理 4 仅肯定限定考虑一个预先给定的集合 X 时可以做到), 而正是由于忽略了这一谨慎考虑, 导致数学家在 19 世纪末发现了著名的 "集合论的悖论".

作为例子, 考虑关系 $x \notin x$. 假定存在一个集合 A 使得关系 $x \in \mathrm{A}$ 和 $x \notin x$ 是等价的; 以 A 代换 x, 利用 §0 的规则 (TL7) 即得关系 $\mathrm{A} \in \mathrm{A}$ 等价于它的否定 $\mathrm{A} \notin \mathrm{A}$, 换句话说, 数学是矛盾的! (说真的, 数学未必没有矛盾, 不过每当出现矛盾, 人们总会得到结论: 导致矛盾的前提是假的.)

所有集合的集合的概念 —— 它涉及的是一个这样的集合 X, 对于所有 x 有 $x \in \mathrm{X}$ —— 也是矛盾的, 因为利用定理 4, 它允许谈论所有使得 $x \notin x$ 的集合, 而刚刚看到, 这是不可能的.

这些例子指出, 在数学中 "集合" 这个词的使用要遵从直观不能表明的限制.

作为定理 4 的例证, 我们取一个集合 X, X 的一个子集 M, 而关系 $\mathrm{R}\{x\}$ 是 $x \notin \mathrm{M}$, 这样得到的 X 的子集 A 称为 M **在** X **里的补集**, 记作

$$\mathrm{X} - \mathrm{M} \quad 或 \quad \complement_{\mathrm{X}}\mathrm{M}$$

(我们仅使用记号 $\mathrm{X} - \mathrm{M}$). 这是 X 的不属于 M 的元素的集合.

定理 5 设 M 和 N 是 X 的子集, 则关系

$$\mathrm{M} \subset \mathrm{N} \quad 和 \quad \mathrm{X} - \mathrm{N} \subset \mathrm{X} - \mathrm{M}$$

是等价的. 对于 X 的所有子集 M, 我们有

$$\mathrm{X} - (\mathrm{X} - \mathrm{M}) = \mathrm{M}.$$

集合 $X-(X-M)$ 由使得关系 $x \in (X-M)$ 是假的 $x \in X$ 所组成; 而对于 $x \in X$ 等价于 $x \in M$ 的否定, 从而集合 $X-(X-M)$ 由使得 $x \in M$ 的否定是假的 $x \in X$ 所组成, 即关系 $x \in M$ 是真的, 由此得到

$$X-(X-M)=M.$$

假定 $M \subset N \subset X$. 由于关系 $x \in M$ 蕴含关系 $x \in N$, 第二个的否定蕴含第一个的否定, 于是关系 $x \in X-N$ 蕴含关系 $x \in X-M$, 从而我们有 $X-N \subset X-M$. 反之, 由于同样的推理, 这个关系蕴含

$$X-(X-M) \subset X-(X-N),$$

即 $M \subset N$.

¶注 6 前面的 "证明" 逻辑上十分不充分, 因为它使用了 "集合" 这个词的直观意义. 比如说, 为了正确证明关系 $X-(X-M)=M$, 应该引进关系



$$R: x \in M, \quad S: x \in X.$$

假设 M 是 X 的一个子集意指 R 蕴含 S, 从而如果 R 蕴含 S, 则关系 $X-(X-M)=M$ 意指 R 和关系

$$S \wedge [\neg(S \wedge (\neg R))]$$

是等价的. 为了建立这两个关系的等价性, 人们自然应该使用 §0 的证明规则.

这告诉我们看起来显然的断言, 当真正实际证明它时, 不再是简单的. 希腊人已经注意到这一事实.

4. 空集

设 X 是一个集合, 在 X 的子集中出现 X 本身, 这就允许考虑集合

$$\varnothing = X-X,$$

称为 X 的**空子集**. 关系 $x \in \varnothing$ 因此等价于关系

$$x \in X \quad \text{和} \quad x \notin X$$

的合取, 从而显然不存在任何对象 $x \in \varnothing$. 集合 $\varnothing = X-X$ 不依赖集合 X, 换句话说, 对于任意集合 X 和 Y, 我们有

$$X-X=Y-Y.$$

事实上, 不难理解 $X-X$ 和 $Y-Y$ 确实具有相同的元素, 因为它们根本不具有任何元素.

¶注 7 上面的证明其实算不上一个证明. 按照定理 2 应当证明关系 $x \in \mathrm{X} - \mathrm{X}$ 和关系 $y \in \mathrm{Y} - \mathrm{Y}$ 是等价的; 用 R 表示关系 $x \in \mathrm{X}$, 而用 S 表示关系 $x \in \mathrm{Y} - \mathrm{Y}$, 所有都归结为证明关系

$$\mathrm{R} \wedge (\neg \mathrm{R})$$

等价于关系

$$\mathrm{S} \wedge (\neg \mathrm{S}),$$

而这来自更一般的事实, 关系 $\mathrm{R} \wedge (\neg \mathrm{R}) \Rightarrow \mathrm{T}$ 是真的 (§0, 习题 2).

集合 $\varnothing = \mathrm{X} - \mathrm{X}$, 因而它总是同一个, 基于这个理由称为**空集**; 它不具有任何元素, 更精确地表示为对于任意 x, $x \in \varnothing$ 是假的.

显然对于所有集合 X, 我们有

$$\varnothing \subset \mathrm{X},$$

并且这个性质刻画了空集的特征. 因为如果两个集合 A 和 B 满足对于所有集合 X 有 $\mathrm{A} \subset \mathrm{X}$ 和 $\mathrm{B} \subset \mathrm{X}$, 则我们发现特别地有 $\mathrm{A} \subset \mathrm{B}$ 和 $\mathrm{B} \subset \mathrm{A}$, 故 $\mathrm{A} = \mathrm{B}$.

5. 一个和两个元素的集合

设 x 是一个数学对象, 那么存在唯一的具有下列性质的一个集合 $\{x\}$: 关系

$$y \in \{x\} \quad \text{等价于} \quad y = x.$$

这种类型的集合称为**一个元素的集合**. 显然关系

$$\{x\} = \{y\}, \quad x = y$$

对于任何 x 和 y 是等价的.

我们十分经常地使用下列结果 (不明确地引用它):

定理 6 一个集合 X 是一个元素的集合, 必须且只需它满足条件

a) X 是非空的;

b) 对于所有 $x \in \mathrm{X}$ 和 $y \in \mathrm{X}$ 有 $x = y$.

条件显然是必要的. 反之, 假定它们成立并选择 X 的一个元素 x (因为 X 是非空的, 这是可能的). 关系 $y = x$ 蕴含 $y \in \mathrm{X}$, 反之根据假设关系 b), $y \in \mathrm{X}$ 蕴含 $y = x$. 条件 $y \in \mathrm{X}$ 和 $y \in \{x\}$ 因而是等价的, 这就证明了 $\mathrm{X} = \{x\}$ 并且完成了证明.

现在设 x 和 y 是两个数学对象; 存在一个且仅仅一个集合, 记作 $\{x, y\}$, 它仅有的元素是 x 和 y, 换句话说, 关系

$$z \in \{x, y\}$$

等价于关系

$$z = x \quad 或 \quad z = y.$$

如果 $x \neq y$, 这种类型的集合称为**两个元素的集合**. 如果 $x = y$, 显然

$$\{x, y\} = \{x, x\} = \{x\}$$

是一个元素的集合.

同样定义三个, 四个, $\cdots$ 元素的集合. 这样得到的集合称为**有限集合**, 其他的集合称为**无穷集合**, 在 §5 将详细研究这两个概念.

¶注 8 一个, 两个, 三个, $\cdots$ 元素集合的存在性直观上是显然的, 不能够证明 —— 更严格地说, 断言 "对于任意元素 x 和 y 存在一个其仅有的元素是 x 和 y 的集合" 是一个公理, 由它可以容易地推导出其他级别的有限集的存在性.

我们补充说明无穷集的存在性也是一个数学公理 (自然数的集合是无穷的, 但是我们并没有给出自然数的数学定义, 也没有证明它们属于同一个集合). 参见 §5.

6. 一个给定集合的子集的集合

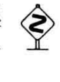
设 X 是一个集合, 存在一个且仅一个集合, 记作

$$\mathscr{P}(\mathrm{X}),$$

具有下列性质: $\mathscr{P}(\mathrm{X})$ 的元素是 X 的子集. 换句话说关系

$$\mathrm{Y} \in \mathscr{P}(\mathrm{X}) \quad 和 \quad \mathrm{Y} \subset \mathrm{X}$$

是等价的, 这是数学公理中的一个. 我们说 $\mathscr{P}(\mathrm{X})$ 是 **X 的子集的集合**.



注 9 从一个集合 X 过渡到集合 $\mathscr{P}(\mathrm{X})$ 的运算容许构造越来越复杂的集合. 后面 (§5) 将看到, 如果 X 是一个 n 个元素组成的有限集合, 则 $\mathscr{P}(\mathrm{X})$ 是有限的并且由 2^n 个元素组成. 于是, 集合

$$\varnothing, \mathscr{P}(\varnothing), \mathscr{P}(\mathscr{P}(\varnothing)), \mathscr{P}(\mathscr{P}(\mathscr{P}(\varnothing))), \mathscr{P}(\mathscr{P}(\mathscr{P}(\mathscr{P}(\varnothing)))), \cdots$$

分别由

$$0,\ 1,\ 2,\ 2^2 = 4,\ 2^4 = 16,\ 2^{16} = 65536,\ 2^{65536}, \cdots$$

个元素组成. 从空集出发, 能够构成如此复杂的迅速变得无法实际计数的集合.

此外还有, 对于所有集合 X 有

$$\mathrm{X} \in \mathscr{P}(\mathrm{X}).$$

正如在注 1 中曾陈述过的那样, 这指出对象的所有集合本身是另外一个集合的一个"元素"—— 明白这一点的最简单的方式是写出关系

$$\mathrm{X} \in \{\mathrm{X}\}.$$

§1 习题

¶¶1. 利用 Hilbert 运算, 空集的完全数学定义是: $\varnothing$ 表示数学对象

$$\tau_{\mathrm{X}}[(\forall x)(x \notin \mathrm{X})];$$

由此推出用形式化语言表达的 $\varnothing$ 的定义 (即以仅含有基本符号的汇集且排除所有缩写的形式写出 $\varnothing$).

¶¶2. 构造定理 5 的完整逻辑证明 (参见注 6).

3. 写出下列集合的所有元素:

$$\mathscr{P}(\mathscr{P}(\mathscr{P}(\mathscr{P}(\varnothing)))).$$

4. 设 X 和 Y 是两个集合, 证明关系 $\mathrm{X} \subset \mathrm{Y}$ 和 $\mathscr{P}(\mathrm{X}) \subset \mathscr{P}(\mathrm{Y})$ 是等价的.

¶5. 证明不存在任何集合 X, 对于它下列关系是真的: $\mathscr{P}(\mathrm{X}) \subset \mathrm{X}$.

6. 设 X 是满足 $0 < x < 1/100000000$ 的数 x 的集合, 而 Y 是满足 $0 < y \leqslant 100000000$ 的数 y 的集合; 证明 $\mathrm{X} \subset \mathrm{Y}$.

§2 函数概念

1. 序偶

我们在前一节已经引进了符号 "=" 和 "∈", 用以构造关系. 现在我们要引进一个运算, 用以构造数学对象.

这个运算在于借助两个数学对象 x 和 y, 按这里书写的次序构成第三个对象, 记作

$$(x, y)$$

称为序偶 (x, y). 构成序偶的运算遵守一个唯一的使用规则: 为了

$$(x, y) = (u, v),$$

成立, 必须且只需

$$x = u \quad 和 \quad y = v.$$

特别的, 仅当 $x = y$ 时有 $(x, y) = (y, x)$, 这说明出现在序偶里的两个对象的前后次序是本质的. 特别要注意不要混淆序偶 (x, y) 和在 §1 里定义的集合 $\{x, y\}$.

注 1 我们可以考虑序偶概念作为一个基本符号, 同符号 

$$\vee, \quad \neg, \quad \tau, \quad \square, \quad =, \quad \in$$

和字母一起, 使得能够用形式化语言写出数学. 但是序偶概念还可以借助其他基本符号 (或归结为基本符号的缩写) 表达: 只需取集合

$$\{\{x\},\{x,y\}\}$$

作为序偶 (x,y) 的定义, 这个集合的元素是集合 $\{x\}$ 和集合 $\{x,y\}$ —— 事实上可以看出, 如此定义的序偶满足给出序偶相等的条件的基本公理. 不过, 这第二个定义把重点放在序偶概念的完全没兴趣的表示上, 因此人们更愿意采取前面第一种表示方法, 事实上对数学而言唯一重要的是知道两个序偶相等的条件.

我们说一个对象 z 是一个序偶, 如果存在对象 x 和 y 使得 $z=(x,y)$. 由于前面陈述的规则, x 和 y 由 z 完全确定. 我们说 x 是 z 的**第一个投影**, 而 y 是 z 的**第二个投影**, 记作

$$x=\mathrm{pr}_1(z), \quad y=\mathrm{pr}_2(z).$$

如果 G 的所有元素是序偶, 则称 G 是一个**图**. 此时, 存在用下列条件刻画其特征的两个集合 X 和 Y: 关系 $x\in \mathrm{X}$ (对应的 $y\in \mathrm{Y}$) 等价于关系: 存在 $z\in \mathrm{G}$ 使得 $x=\mathrm{pr}_1(z)$ (对应的 $y=\mathrm{pr}_2(z)$). 这记作

$$\mathrm{X}=\mathrm{pr}_1(\mathrm{G}), \quad \mathrm{Y}=\mathrm{pr}_2(\mathrm{G}).$$

可以如下扩充序偶概念. 给定三个对象 x, y, z, 令

$$(x,y,z)=((x,y),z),$$

我们称 (x,y,z) 是一个**三元有序组**. 为了有

$$(x',y',z')=(x'',y'',z'')$$

必须且只需

$$x'=x'', \quad y'=y'', \quad z'=z''.$$

事实上, 所考虑的关系写出来就是 $((x',y'),z')=((x'',y''),z'')$, 于是这等价于 $(x',y')=(x'',y'')$ 和 $z'=z''$, 随之等价于 $x'=x''$, $y'=y''$ 和 $z'=z''$.

同样, 给定四个对象 x,y,z,t, 令

$$(x,y,z,t)=((x,y,z),t),$$

我们称 (x,y,z,t) 是一个**四元有序组**, 以此类推.

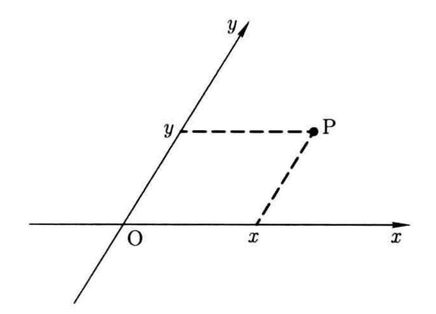
2. 两个集合的笛卡儿乘积

设 X 和 Y 是两个集合, (借助 §0 的方法) 可以证明存在一个以下列性质刻画其特征的集合 Z: 为了有 $z \in \mathrm{Z}$, 必须且只需存在 $x \in \mathrm{X}$ 和 $y \in \mathrm{Y}$, 使得 $z = (x, y)$. 我们说 Z 是 X 和 Y 的**笛卡儿乘积**, 并且书写为

$$\mathrm{Z} = \mathrm{X} \times \mathrm{Y}.$$

因此乘积 $\mathrm{X} \times \mathrm{Y}$ 这个集合是序偶 (x, y) 的集合, 其中 $x \in \mathrm{X}$ 并且 $y \in \mathrm{Y}$.

前述运算以笛卡儿命名的解释如下. 在初等几何的“平面”上, 取两个坐标轴 Ox 和 Oy 和这些轴上的长度单位, 这时就可以定义平面的所有的点的横坐标和纵坐标 x 和 y, 不区分点 P 和序偶 (x, y) 是自然的. 此外, 如果 P′ 是坐标为 (x', y') 的点, 关系 $\mathrm{P} = \mathrm{P}'$ 显然等价于 $x = x'$ 和 $y = y'$, 即 $(x, y) = (x', y')$. 用 $\mathbf{R}$ 表示实数集合, 我们看到平面的一个坐标系的选取就使得平面类似于实数的序偶的集合 $\mathbf{R} \times \mathbf{R} = \mathbf{R}^2$.



设 A, B, X, Y 是四个集合, 则关系

$$\mathrm{A} \subset \mathrm{X} \text{ 和 } \mathrm{B} \subset \mathrm{Y} \quad \text{蕴含} \quad \mathrm{A} \times \mathrm{B} \subset \mathrm{X} \times \mathrm{Y}.$$

只要 A 和 B 都不是空集, 这个断言的逆是正确的, 如果 $\mathrm{A} \times \mathrm{B} \subset \mathrm{X} \times \mathrm{Y}$, 并且 B 至少含有一个元素 b, 那么对于所有 $a \in \mathrm{A}$ 我们有 $(a, b) \in \mathrm{A} \times \mathrm{B}$, 因此 $(a, b) \in \mathrm{X} \times \mathrm{Y}$, 随之对于一个 $x \in \mathrm{X}$ 和一个 $y \in \mathrm{Y}$ 有 $(a, b) = (x, y)$, 于是对于一个 $x \in \mathrm{X}$ 有 $a = x$, 故 $\mathrm{A} \subset \mathrm{X}$. 同样可以证明如果 A 是非空的, 则有 $\mathrm{B} \subset \mathrm{Y}$.

如果集合 A, B 当中有一个是空集, 那么总有 $\mathrm{A} \times \mathrm{B} \subset \mathrm{X} \times \mathrm{Y}$, 理由是简单的,

$$\text{如果 } \mathrm{A} = \varnothing \text{ 或 } \mathrm{B} = \varnothing, \text{则 } \mathrm{A} \times \mathrm{B} = \varnothing.$$

事实上, 如果 $\mathrm{A} \times \mathrm{B} = \varnothing$ 是假的, 则 $\mathrm{A} \times \mathrm{B} = \varnothing$ 至少含有一个序偶 (x, y), 从而 $x \in \mathrm{A}$, 并且 $y \in \mathrm{B}$, 这就证明了 A 和 B 都不是空集.

笛卡儿乘积的概念可以推广到多个集合的情形. 设 $\mathrm{X}, \mathrm{Y}, \mathrm{Z}, \mathrm{T}, \cdots$ 是集合, 定义

$$\mathrm{X} \times \mathrm{Y} \times \mathrm{Z} = (\mathrm{X} \times \mathrm{Y}) \times \mathrm{Z}; \quad \mathrm{X} \times \mathrm{Y} \times \mathrm{Z} \times \mathrm{T} = (\mathrm{X} \times \mathrm{Y} \times \mathrm{Z}) \times \mathrm{T}; \quad \cdots$$

$\mathrm{X} \times \mathrm{Y} \times \mathrm{Z}$ 的元素显然是在第 1 小节中定义的三元有序组, 其中 $x \in \mathrm{X}, y \in \mathrm{Y}, z \in \mathrm{Z}$. 同样 $\mathrm{X} \times \mathrm{Y} \times \mathrm{Z} \times \mathrm{T}$ 的元素显然是四元有序组, 其中 $x \in \mathrm{X}, y \in \mathrm{Y}, z \in \mathrm{Z}, t \in \mathrm{T}$.

注意关系

$$(\mathrm{X} \times \mathrm{Y}) \times \mathrm{Z} = \mathrm{X} \times (\mathrm{Y} \times \mathrm{Z})$$

是假的. 事实上左端的元素是形如 $((x, y), z)$ 的对象, 而右端的元素是形如 $(x, (y, z))$ 的对象, 两个序偶相等的规则不允许写出 $((x, y), z) = (x, (y, z))$, 不管 x, y, z 是什么. 但是, 在实际中, 人们约定对 $((x, y), z)$ 和 $(x, (y, z))$ 不做任何区别, 并且认为 $(\mathrm{X} \times \mathrm{Y}) \times \mathrm{Z}$ 和 $\mathrm{X} \times (\mathrm{Y} \times \mathrm{Z})$ 是相等的. 对于语言的这个约定严格地说是矛盾的, 正如后面我们将引入的其他约定一样, 不过使用这个约定导致的矛盾是 "无关紧要的". 读者在逐渐习惯了集合论的一些推理之后, 就能够避免陷入这种困惑之中.

最后, 如果 X 是一个集合, 则令

$$\mathrm{X}^2 = \mathrm{X} \times \mathrm{X}, \quad \mathrm{X}^3 = \mathrm{X} \times \mathrm{X} \times \mathrm{X}, \quad \mathrm{X}^4 = \mathrm{X} \times \mathrm{X} \times \mathrm{X} \times \mathrm{X},$$

以此类推. 例如, 如果 $\mathbf{R}$ 表示实数的集合, $\mathbf{R}^4$ 就是由四个实数组成的四元有序组 (x, y, z, t) 的集合, 这就是物理学家所谓的 "四维空间" 或 "空间—时间". 初学者不要被这个专门术语吓倒, 经验证明在 $\mathbf{R}^4$ 里进行推理不会比在 $\mathbf{R}^2$ 或 $\mathbf{R}^{100}$ 里更困难.

显而易见, 跟两个集合的乘积的情况一样, 集合的乘积是空集, 只要一个因子是空集.

3. 图像和函数

设 X 和 Y 是两个集合, 称令 X 的每一个元素 x 对应 Y 的一个按一个确定的规律依赖 x 的元素 y 的运算为定义在集合 X 上在 Y 内取值的函数. 例如当 $\mathrm{X} = \mathrm{Y} = \mathbf{R}$ 时的函数 $y = \sin x$.

不幸的是, 按原文照抄的这个所谓定义数学上包含许多数学上未曾定义过的词汇, 例如, "令 …… 对应" 意指什么? 轻信这个定义, 就会又一次仅仅得到一个文字游戏.

故被迫修改这个定义, 用下面的定义代替它 (在经典情形, 这定义回归到由其曲线表示给定的函数, 物理学家认为这是极其合理的方法): 给定集合 G, X, Y, 称一个满足下列条件的一个三元有序组

$$f = (\mathrm{G}, \mathrm{X}, \mathrm{Y})$$

为函数:

(F1) $\mathrm{G} \subset \mathrm{X} \times \mathrm{Y}$;

(F2) 对于每个 $x \in \mathrm{X}$, 存在一个且仅一个 $y \in \mathrm{Y}$ 使得 $(x, y) \in \mathrm{G}$.

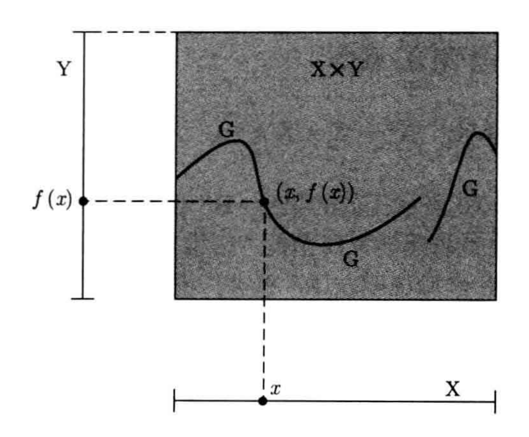
条件 (F1) 意即 G 是一个图 (第 1 小节), 我们称 G 是**函数 f 的图**. 根据 (F2), 对于每个 $x \in \mathrm{X}$ 存在一个 $z \in \mathrm{G}$ 使得 $x = \mathrm{pr}_1(z)$, 于是有

$$\mathrm{pr}_1(\mathrm{G}) = \mathrm{X}, \quad \mathrm{pr}_2(\mathrm{G}) \subset \mathrm{Y}.$$

如果 x 是 X 的一个元素, Y 的唯一使得 $(x, y) \in \mathrm{G}$ 的元素 y 称为**函数 f 在 x 的值**, 用记号

$$y = f(x)$$

表示它. 显然 f 的图 G 是形如 $(x, f(x))$ 的序偶的集合, 其中 $x \in \mathrm{X}$, 这跟人们对于函数的直观想法是符合的.



给定函数 $f = (\mathrm{G}, \mathrm{X}, \mathrm{Y})$, 我们称 X 是 f 的**出发集**, 而 Y 是 f 的**到达集**.

给定两个集合 X 和 Y, 称所有以 X 为出发集以 Y 为到达集的函数为**从 X 到 Y 内的映射**. 词汇 "函数" 和 "映射" 是同义词, 但是在实际中, 说 "f 是一个从 X 到 Y 内的映射" 比说 "f 是一个定义在 X 上在 Y 内取值的映射" 方便. 此外, 代替说

f 是一个从 X 到 Y 内的映射,

经常说

给定一个映射 $f : \mathrm{X} \to \mathrm{Y}$,

或

给定一个映射 $\mathrm{X} \xrightarrow{f} \mathrm{Y}$.

还可能出现这种情形, 代替 f, g 等这样的字母而用 "公式" 表示函数, 用公式根据 x 计算 $f(x)$; 这样, 在 $\mathrm{X} = \mathrm{Y} = \mathbf{R}$ 的情形, 当我们说

考虑从 $\mathbf{R}$ 到 $\mathbf{R}$ 的映射 $x \to x^3$

或当我们说

$$\text{考虑 } \mathbf{R} \text{ 上的函数 } x^3,$$

应当说成

$$\text{考虑从 } \mathbf{R} \text{ 到 } \mathbf{R} \text{ 的映射 } f, \text{ 使得对于每个 } x \in \mathbf{R} \text{ 有 } f(x) = x^3,$$

甚至应该这样说:

$$\text{考虑映射 } f = (\mathrm{G}, \mathbf{R}, \mathbf{R}), \text{ 其中 G 是满足 } y = x^3 \text{ 的序偶 } (x, y) \in \mathbf{R} \times \mathbf{R} \text{ 的集合.}$$

给定两个函数

$$f = (\mathrm{G}, \mathrm{X}, \mathrm{Y}) \quad \text{和} \quad f' = (\mathrm{G}', \mathrm{X}', \mathrm{Y}')$$

考虑到两个三元有序组相等的条件, 等式 $f = f'$ 表示我们有

$$\mathrm{G} = \mathrm{G}', \quad \mathrm{X} = \mathrm{X}', \quad \mathrm{Y} = \mathrm{Y}'.$$

由于 $\mathrm{X} = \mathrm{pr}_1(\mathrm{G})$ 和 $\mathrm{X}' = \mathrm{pr}_1(\mathrm{G}')$, 第一个条件蕴含第二个. 此外, 如果条件

$$\mathrm{X} = \mathrm{X}', \quad \mathrm{Y} = \mathrm{Y}'$$

已经验证, 为了有 $\mathrm{G} = \mathrm{G}'$ (即 $f = f'$), 显然只需

$$\text{对于所有 } x \in \mathrm{X},\ f(x) = f'(x).$$

注 2 由于初学者知道的例子不多, 他们常常就把所有函数设想为由通过 x 计算 $f(x)$ 的公式定义的. 关于函数概念的最错误的想法莫过于此.



首先只要没有给出精确定义, “公式” 这个词就什么也没有说. 似乎是这样, 对于初学者来说, 一个 “公式” 就是施于变量 x 的一系列代数运算. 数学上, 当 x 是一个任意集合的元素时, 这类运算没有任何意义. 反之如果变量 x 是一个实数, 对于它事实上可以实行代数运算, 用这种方式得到的函数是如此特殊, 以致摒弃这种定义函数概念的想法少说也有两百年了. 事实上, 分析的需要逼迫数学家发现越来越一般的函数, 而研究这类或多或少 “任意的” 函数导致集合论的发现, 随之是积分的现代理论, 拓扑空间理论, 等等. 换句话说, 随着函数概念的不断推广, 指引人们构筑了现代数学的一个美好部分.

人们对于 “代数公式” 定义的函数的研究并不缺乏兴趣, 反之这是数学的一个分支 (代数几何) 的研究对象, 它从没有像今日这样活跃. 但是人们用以研究这些函数的方法以及在其研究中所提出的课题, 跟初等代数和分析所感兴趣的, 没有任何共同之处.

注 3 设 X 和 Y 是两个集合, 那么从 X 到 Y 内的所有映射的集合记为

$$\mathrm{Y}^{\mathrm{X}},$$

并称为**从 X 到 Y 内的映射的集合**. 由于一个函数是一个三元有序组 $f = (\mathrm{G}, \mathrm{X}, \mathrm{Y})$, 其中 $\mathrm{G} \subset \mathrm{X} \times \mathrm{Y}$, 我们看到从 X 到 Y 内的映射的集合包含于 $\mathscr{P}(\mathrm{X} \times \mathrm{Y}) \times \{\mathrm{X}\} \times \{\mathrm{Y}\}$ 内.

在实际中, 当 X 和 Y 给定时, 人们一般把一个从 X 到 Y 内的映射等同于它的图像 $\mathrm{G} \subset \mathrm{X} \times \mathrm{Y}$. 按照这个约定, 从 X 到 Y 内的映射的集合作为集合 $\mathscr{P}(\mathrm{X} \times \mathrm{Y})$ 的一个子集而出现.

注 4 经常使用一个跟函数概念临近的**族**的概念, 其直观定义是: 给定一个集合 I, 为了构造一个**以 I 为指标集的族** (或**以 I 标记的族**), 对于每个 $i \in \mathrm{I}$ 给定一个依赖 i 的对象 (没有预先明确在哪个集合里选取对应于 I 的元素的对象). 如果记 x_i 为对应于 $i \in \mathrm{I}$ 的对象, 则一般用记号

$$(x_i)_{i \in \mathrm{I}}$$

表示所考虑的族. 数学上, 以 I 标记的族是具有下面两个性质的一个图: 我们有 $\mathrm{pr}_1(\mathrm{G}) = \mathrm{I}$, 并且对于每个 $i \in \mathrm{I}$ 存在一个唯一的 $z \in \mathrm{G}$ 使得 $\mathrm{pr}_1(z) = i$; 书写出 $z = (i, x_i)$ 就回到了上面叙述的直观定义.

设 $(x_i)_{i \in \mathrm{I}}$ 是一个族, 如果对于所有 $i \in \mathrm{I}$ 有 $x_i \in \mathrm{X}$, 我们说这是**集合 X 的元素的一个族**. 总存在一个这样的集合 X, 比如 $\mathrm{pr}_2(\mathrm{G})$, 其中 G 是所考虑的族的图. 同样我们说一个族 $(\mathrm{A}_i)_{i \in \mathrm{I}}$ 是**集合 X 的子集的一个族**, 如果对于每个 $i \in \mathrm{I}$ 有 $\mathrm{A}_i \subset \mathrm{X}$.

当一个族的指标集的元素为 $1, 2, 3, \cdots$ 这些自然数时, 就说这个族是一个**序列**, 经常用记号

$$(x_n)_{n \geqslant 1}$$

表示一个序列. 给定一个集合 X 的元素的一个序列归结为选择 X 的元素

$$x_1, x_2, \cdots, x_n, \cdots$$

或给定从自然数的集合到 X 内的一个映射 $n \to x_n$.

不言而喻, 我们承认一个集合 (记为 $\mathbf{N}$) 的存在性, 其元素是数 $0, 1, \cdots$ 这个集合的存在性是显然的, 只要给出 "集合" 和 "自然数" 这两个词的直观意义. 至于数学上证明 $\mathbf{N}$ 的存在性, 则完全是另一回事, 因为不仅首先要创立一个整数的正确的数学理论, 而且还要承认由无穷个元素组成的集合的存在性; 而此类集合的存在性尽管直观上是显然的, 但不能够证明 —— 这是数学的公理之一.

4. 像和逆像

设 f 是从集合 X 到集合 Y 内的一个映射. 给定 X 的一个子集 A, 称具有下列性质

$$\text{存在 } x \in \mathrm{X} \text{ 使得 } y = f(x)$$

的 $y \in \mathrm{Y}$ 的集合为**在 f 下 A 的像**. 如果 A 缩减为一个单一元素 x, 它的像显然缩减为一个单一元素 $f(x)$. 当 A 是 X 的任意子集时, 用 $f(\mathrm{A})$ 表示 A 的像 —— 这个记号欠妥, 因为严格说来 $f(\mathrm{A})$ 仅对于 $\mathrm{A} \in \mathrm{X}$ 才有意义.

显然对于所有映射 f 有

$$f(\varnothing) = \varnothing.$$

仍然设 f 是从集合 X 到集合 Y 内的一个映射, 而考虑 Y 的一个子集 B, 称使得 $f(x) \in \mathrm{B}$ 的 $x \in \mathrm{X}$ 的集合为**在 f 下 B 的逆像**, 记作

$$\overset{-1}{f}(\mathrm{B}),$$

经常 (不恰当地) 写成 $f^{-1}(\mathrm{B})$.

显然成立关系:

对于 X 的所有子集 A, 有 $\mathrm{A} \subset \overset{-1}{f}(f(\mathrm{A}))$,

对于 Y 的所有子集 B, 有 $\mathrm{B} \supset f(\overset{-1}{f}\mathrm{B}))$;

但是我们不能在这些关系中把包含换做相等.

我们称映射 $f: \mathrm{X} \to \mathrm{Y}$ 在 X **的子集 A 上是常值的**, 如果 $f(\mathrm{A})$ 缩减为一个单一的元素, 即 (§1, 定理 6) 如果

$$\text{对于任意 } x' \in \mathrm{A} \text{ 和 } x'' \in \mathrm{A} \text{ 有 } f(x') = f(x'').$$

我们称 f 是**常值的**, 如果 f 在整个 X 上是常值的.

设 f 是从 X 到自身内的一个映射, 我们称 X 的一个子集 A **在 f 下是稳定的**, 如果 $f(\mathrm{A}) \subset \mathrm{A}$, 换句话说, 如果 $x \in \mathrm{A}$ 蕴含 $f(x) \in \mathrm{A}$. 当 A 缩减为一个单一元素 x 时, 这意味着

$$f(x) = x,$$

我们称这样的 x 是 f 的一个**不动点**.

一个集合在一个映射下的像的概念会出现在初等几何里, 比如我们说 "一条直线在旋转下的变换", 那么 "一条直线" 是 (点的) 一个集合, "一个旋转"是 (从空间的点的集合到空间自身内的) 某一个映射, 而所说的 "变换" 正是所考虑的集合在这个映射下的像.

5. 函数的限制和延拓

设 $f = (\mathrm{G}, \mathrm{X}, \mathrm{Y})$ 是一个函数, 考虑一个集合 $\mathrm{X}' \subset \mathrm{X}$. 设 G' 是使得 $\mathrm{pr}_1(z) \in \mathrm{G}'$ 的 $z \in \mathrm{G}$ 的集合, 那么 $f' = (\mathrm{G}', \mathrm{X}', \mathrm{Y})$ 是一个函数. 事实上显然 $\mathrm{G}' \subset \mathrm{X}' \times \mathrm{Y}$, 并且对于所有 $x \in \mathrm{X}'$, 存在一个且仅一个 $z = (x, f(x)) \in \mathrm{G}'$ 使得 $\mathrm{pr}_1(z) = x$. 从 X' 到 Y 的使得

$$\text{对于所有 } x \in \mathrm{X}' \text{ 有 } f'(x) = f(x)$$

的映射 f' 称为 f **在 X' 上的限制**.



注 5 f 在 X' 上的限制经常记作

$$f' = f|X' \quad \text{或} \quad f_{X'}.$$

此外, 设给定两个映射 f 和 g, 它们的出发集包含一个同样的集合 X, 如果

$$\text{对于所有 } x \in X \text{ 有 } f(x) = g(x),$$

则称 **f 和 g 在 X 上重合**. 举例说, 如果 $f|X = g|X$, 则 f 和 g 在 X 上重合. 但反之不真, 虽然看起来似乎是真的.

设 $f = (G, X, Y)$ 和 $f' = (G', X', Y')$ 是两个映射, 如果

$$X' \subset X,\ Y' \subset Y, \text{ 并且对于所有 } x \in X' \text{ 有 } f(x) = f'(x),$$

则称 **f 是 f' 的一个延拓**. 比如, 如果取 f' 为 f 在 X 的一个子集上的限制, 则就是这种情形. 注意下列结果: *设 X', X, Y 是三个集合, $X' \subset X$, 而 f 是从 X' 到 Y 内的一个映射. 如果 Y 是非空的, 则存在从 X 到 Y 的延拓 f 的一个映射.* 为了构造一个延拓 f 的映射 $g : X \to Y$, 选定 Y 的一个元素 c, 令

$$g(x) = \begin{cases} f(x), & \text{如果 } x \in X', \\ c, & \text{如果 } x \notin X'. \end{cases}$$

显然存在其他延拓 f 的方式, 这里给出的是其中最简单的一个.

6. 复合映射

设 X, Y, Z 是三个集合, 而

$$f = (G, X, Y), \qquad g = (H, Y, Z)$$

分别是从 X 到 Y 内和从 Y 到内 Z 的两个映射. 我们如下定义从 X 到 Z 内的第三个映射

$$h = (K, X, Z),$$

$$h(x) = g(f(x)), \text{ 对于所有 } x \in X,$$

h 的图显然是具有下列性质的序偶 (x, z) 的集合: 存在 $y \in Y$, 使得 $(x, y) \in G$, 并且 $(y, z) \in H$.

映射 h 用记号

$$h = g \circ f$$

表示, 称为 **f 和 g 的复合**. 仅当 f 的到达集等于 g 的出发集时, 复合 $g \circ f$ 才有定义.

注 6 复合映射的概念如今代替了往日在一些特殊情形使用的 "函数的函数" 和 "两个变换的乘积" 的概念. 比如取实数集 $\mathrm{X} = \mathrm{Y} = \mathrm{Z} = \mathbf{R}$, $f(x) = x^2$, $g(x) = \sin x$, 则 $g \circ f$ 是函数

$$x \to \sin(x^2),$$

而 $f \circ g$ 是函数

$$x \to \sin^2 x,$$

这表明一般 $f \circ g \neq g \circ f$ (当两端有定义时). 当 $\mathrm{X} = \mathrm{Y} = \mathrm{Z} =$ 空间 (在 "空间几何" 的意义下 —— 我们不打算在这里给出正确的定义) 时, 可以取 f 和 g 是在几何意义下的变换 —— 旋转, 平移, 相似, 等等; 复合映射 $f \circ g$ 就是初等几何里的变换 f 和 g 的 "乘积".

我们现在证明一个定理, 它在后面会起到重要的作用.

定理 1 *设 $\mathrm{X}, \mathrm{Y}, \mathrm{Z}$ 是三个集合, 给定两个映射*

$$f : \mathrm{X} \to \mathrm{Y}, \quad h : \mathrm{X} \to \mathrm{Z},$$

则以下条件是等价的:

a) 存在一个映射

$$g : \mathrm{Y} \to \mathrm{Z}$$

满足 $h = g \circ f$;

b) 对于任意 $x', x'' \in \mathrm{X}$, 关系

$$f(x') = f(x'')$$

蕴含关系

$$h(x') = h(x'').$$

条件 a) 蕴含条件 b), 因为如果 a) 满足, 则有

$$h(x') = g(f(x')) = g(f(x'')) = h(x'').$$

我们现在要证明 b) 蕴含 a).

首先考察特殊情形: $f(\mathrm{X}) = \mathrm{Y}$. 为了构造 g, 我们要如下构造其图像 $\mathrm{G} \subset \mathrm{Y} \times \mathrm{Z}$: G 是这样的序偶 (y, z) 的集合, 至少存在一个 $x \in \mathrm{X}$ 满足

$$y = f(x), \quad z = h(x).$$

[G 的这个结构是自然的, 因为如果假定问题已经解决, 而 G 是序偶 $(y, g(y))$ 的集合; 由于 $f(\mathrm{X}) = \mathrm{Y}$, 可以写出 $y = f(x)$, 于是有 $g(y) = g(f(x)) = h(x)$, 从而 G 必然由形

如 $(f(x), h(x))$ 的序偶组成, 其中 x 遍历 X.] 我们来验证如此得到的 G 事实上是一个映射 $g: \mathrm{Y} \to \mathrm{Z}$ 的图像, 并且满足 $h = g \circ f$. 为了证明 G 是一个映射的图像, 应当验证对于所有 $y \in \mathrm{Y}$, 存在唯一的一个 $z \in \mathrm{Z}$ 使得 $(y, z) \in \mathrm{G}$. 至少存在一个这样的 z 是显然的: 只需选择一个 x 使得 $y = f(x)$, 再取 $z = h(x)$; 此外, 假定 G 含有 (y, z') 和 (y, z''), 那么在 X 内存在元素 x' 和 x'' 使得

$$y = f(x'), \quad z' = h(x'),$$
$$y = f(x''), \quad z'' = h(x'').$$

于是 $f(x') = f(x'')$, 再根据假设 b), $h(x') = h(x'')$, 即 $z' = z''$, 这就指出 G 是一个从 Y 到 Z 内的一个映射 g 的图像. 为了证明 $h = g \circ f$, 考虑一个 $x \in \mathrm{X}$. 由 G 的构造, 它含有序偶 $(f(x), h(x))$, 从而 $h(x) = g(f(x))$, 这就证明了所要的关系.

接下来要证明在一般情形下 b) 蕴含 a). 令 $\mathrm{Y}' = f(\mathrm{X})$, 并且考虑从 X 到 Y′ 的映射 f': 对于所有 $x \in \mathrm{X}$, 令 $f'(x) = f(x)$. 用 $\mathrm{X}, \mathrm{Y}', \mathrm{Z}, f', h$ 代替 $\mathrm{X}, \mathrm{Y}, \mathrm{Z}, f, h$, 那么现在就处于 b) 的假设之下, 并且 $f'(\mathrm{X}) = \mathrm{Y}'$. 根据刚刚证明的结果, 存在一个从 Y′ 到 Z 内的一个映射 g' 使得 $h = g' \circ f'$, 延拓 g' 成一个从 Y 到 Z 内的一个映射 g (在第 5 小节已经看到这是可能的). 对于所有 $x \in \mathrm{X}$ 有

$$h(x) = g'(f'(x)) = g(f'(x)) = g(f(x)),$$

故 $h = g \circ f$, 这就完成了证明.

定理 1 的 "直观" 解释如下: 由于关系 $f(x') = f(x'')$ 蕴含关系 $h(x') = h(x'')$, 为了计算 $h(x)$, 只需知道 $f(x)$, 于是 $h(x)$ 就应当是 $f(x)$ 的一个函数. 但显然这种解释算不上一个正确的证明.

映射的复合运算的最重要的性质是下列定理表达的 "结合性":

定理 2 给定任意的映射

$$f: \mathrm{X} \to \mathrm{Y}, \quad g: \mathrm{Y} \to \mathrm{Z}, \quad h: \mathrm{Z} \to \mathrm{T},$$

则有关系

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

令 $g \circ f = u$ 和 $h \circ g = v$. 在 $x \in \mathrm{X}$, 要证明的关系的左端的值是 $h(u(x)) = h(g(f(x)))$, 而右端的值是 $v(f(x)) = h(g(f(x)))$, 由此显然得到定理.

定理 2 使得能够定义 $h \circ g \circ f$, 以及更一般的表达式

$$f_1 \circ f_2 \circ \cdots \circ f_n,$$

而不致引起混淆, 只要它们有意义, 即每个因子的出发集是随后的因子的到达集. 特别的, 对于所有映射 $u: \mathrm{X} \to \mathrm{X}$ 和所有整数 $q \geqslant 1$, 可以定义

$$u^q = u \circ u \circ \cdots \circ u \ (q \text{ 个因子}).$$

7. 单射

我们说映射 $f: \mathrm{X} \to \mathrm{X}$ 是**单射的**, 如果对于任意 $x', x'' \in \mathrm{X}$, 关系

$$f(x') = f(x'') \quad \text{蕴含 } x' = x'',$$

或

$$x' \neq x'' \quad \text{蕴含 } f(x') \neq f(x'').$$

这样的映射称为**单射**. 代替词汇 "单射的", 过去使用 (大约直到 1955) 形容词**双单值的**, 至今在许多著作中还出现, 这个词现代使用起来不太方便, 原因是没有对应的名词.

例 1 取 $\mathrm{X} = \mathrm{Y} = \mathbf{R}$ 和 $f(x) = x^3$, 那么 f 是单射的, 因为如果 x 和 y 是实数, 关系 $x^3 = y^3$ 蕴含 $x = y$. 反之, 函数 $f(x) = x^2$ 不是单射的, 比如有 $f(-1) = f(1)$.

例 2 对于所有集合 X, 定义**从 X 到 X 内的恒等映射**, 对于每个 $x \in \mathrm{X}$, 令它对应 x 自己, 记为

$$j_{\mathrm{X}},$$

于是有

$$\text{对于所有 } x \in \mathrm{X}, j_{\mathrm{X}}(x) = x.$$

经常用记号 id 代替 j_{X}. 根据定义, 从 X 到 X 内的恒等映射显然是单射的.

我们注意在 $\mathrm{X} \times \mathrm{X}$ 里 j_{X} 的图像是序偶 (x, x) 的集合, 其中的 $x \in \mathrm{X}$, 这个集合称为**乘积 $\mathrm{X} \times \mathrm{X}$ 的对角集**.

当 $\mathrm{X} = \mathbf{R}$, 恒等映射化为中学生熟悉的 "函数 x". 当 X 是平面或空间的点的集合时, 恒等映射只不过是初等几何的 "单位变换".

例 3 设 X 和 Y 是两个集合, 满足 $\mathrm{X} \subset \mathrm{Y}$. 对于所有 $x \in \mathrm{X}$, 令 $j(x) = x$, 这样得到的映射 $j: \mathrm{X} \to \mathrm{Y}$ 称为**从 X 到 Y 内的典范单射** (一般有许多从 X 到 Y 内的单射; 形容词 "典范的" 这里用来表示这个特殊的单射是由 "自然的" 过程得到的, 它不牵扯任何特定元素的选择, 而只涉及它内在的已知条件, 即 Y 和 Y 的一个子集 X).

定理 3 *设 X 和 Y 是非空集合, 而 f 是从 X 到 Y 内的一个映射. 下列性质是等价的:*

a) *f 是单射的;*

b) *存在一个映射 $g: \mathrm{Y} \to \mathrm{X}$, 使得 $g \circ f$ 是从 X 到 X 内的恒等映射.*

事实上, 考虑两个映射

$$f: \mathrm{X} \to \mathrm{Y}, \quad j_{\mathrm{X}}: \mathrm{X} \to \mathrm{X},$$

我们要找一个必要且充分的条件使得存在一个映射 $g:\mathrm{Y}\to\mathrm{X}$, 使得 $j_{\mathrm{X}}=g\circ f$. 这个条件由第 6 小节的定理 1 提供, 即关系 $f(x')=f(x'')$ 蕴含关系 $j_{\mathrm{X}}(x')=j_{\mathrm{X}}(x'')$, 亦即 $x'=x''$, 它恰好表明 f 是单射的; 定理得证.

例 4 取 $\mathrm{X}=\mathbf{R}_+$ 为实数 $x\geqslant 0$ 的集合, 而 $\mathrm{Y}=\mathbf{R}$ 是 (任意符号的) 所有实数的集合; 取 $f(x)=x^2$, 这个映射显然是单射的 (如果用 $\mathbf{R}$ 代替 $\mathbf{R}_+$ 将不是单射的), 故存在一个映射 $g:\mathbf{R}\to\mathbf{R}_+$ 使 $g\circ f$ 为恒等映射, 即使得对于所有 $x\in\mathbf{R}_+$ 有 $g(x^2)=x$, 这个条件显然表示为

$$g(y)=\sqrt{y},\quad \text{如果 } y\geqslant 0.$$

换句话说, 可以取任意一个其值为非负, 当 $y\geqslant 0$ 时与函数 $\sqrt{y}$ 重合的所有函数为 g.

我们注意对于定理 3 显然可以直接证明而不援引定理 1. 要找的映射 g 应当满足关系 $g(f(x))=x$, 由于这是 g 应当满足的唯一条件, 当 y 不属于 $f(\mathrm{X})$ 时可以任意选择 $g(y)$; 反之, 如果 $y\in f(\mathrm{X})$, 因为 f 是单射的存在一个且仅一个 $x\in\mathrm{X}$ 使得 $y=f(x)$, 于是我们应该选择 $g(y)=x$.

8. 满射和双射

如果 $f(\mathrm{X})=\mathrm{Y}$, 换句话说, 如果对于所有 $y\in\mathrm{Y}$ 至少存在一个 $x\in\mathrm{X}$ 使得 $y=f(x)$, 则称映射 $f:\mathrm{X}\to\mathrm{Y}$ 是**满射的**. 反之, 称 f 是单射的, 意指对于所有 $y\in\mathrm{Y}$ 至多存在一个 $x\in\mathrm{X}$ 使得 $y=f(x)$.

我们称一个映射 $f:\mathrm{X}\to\mathrm{Y}$ 是**双射的**, 如果它既是单射的又是满射的, 换句话说, 如果对于所有 $y\in\mathrm{Y}$ 存在一个且仅一个 $x\in\mathrm{X}$ 使得 $y=f(x)$.

例如, 对于所有集合 X, 恒等映射 j_{X} 显然是双射的.

一个满射的映射称为**满射**, 一个双射的映射称为**双射**. 从一个集合 X 到其自身内的双射称为 X 的**置换**. X 的置换的集合用记号

$$\mathfrak{S}(\mathrm{X})$$

表示. 如果 $\mathrm{X}=\{1,2,\cdots,n\}$, 则用 $\mathfrak{S}_n$ 代替 $\mathfrak{S}(\mathrm{X})$.



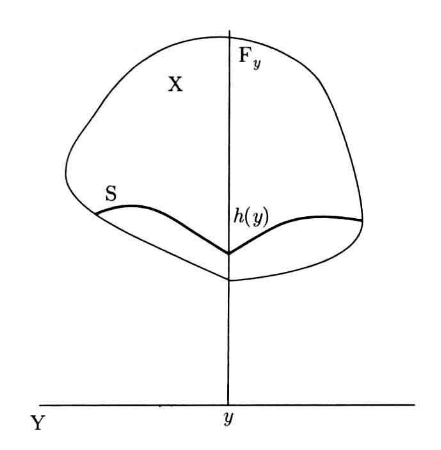
注 7 在早先的术语中, 我们说 "f 是从 X 到 Y 上的一个映射" 以代替说 "f 是从 X 到 Y 内的一个满射", 而为表示 f 是一个双射, 人们经常说 "一一的和映上的".

定理 4 *设 $f:\mathrm{X}\to\mathrm{Y}$ 是一个映射. 以下条件是等价的:*

a) f 是满射的;

b) 存在一个映射 $h:\mathrm{Y}\to\mathrm{X}$ 使得 $f\circ h$ 是从 Y 到 Y 上的恒等映射.

如果 b) 满足, 则对于所有 $y\in\mathrm{Y}$ 有 $f(h(y))=y$, 于是所有 $y\in\mathrm{Y}$ 都有形式 $f(x)$, 即 f 是满射的.



反之假定 f 是满射的, 对于每个 $y \in \mathrm{Y}$, 用 F_y 表示使得 $f(x)=y$ 的 $x \in \mathrm{X}$ 的集合, 这是 X 的非空子集. 为了构造所要找的映射 h, 只需在每个集合 F_y 里随意选取一个元素, 记为 $h(y)$, 我们就定义了从 Y 到 X 内的一个映射 h, 使得对于所有 $y \in \mathrm{Y}$ 有 $f(h(y))=y$, 这就完成了证明.

¶¶注 8 上面证明的第二部分似乎正确无疑, 甚至是显而易见的; 但是数学上看它是远非完善的 (人们难以想象一个证明机在 F_y 里 "随意选取" 一个元素 ……). 正确的证明借助 §0 第 9 小节的 Hilbert 运算而得到: 如果没有更好的选择, 令



$$h(y)=\tau_x(f(x)=y),$$

我们就得到一个函数 $h(y)$.

注意问题归结为构造 X 的一个子集 S, 使得对于每个 $y \in \mathrm{Y}$, 交集 $S \cap \mathrm{F}_y$ 是恰好有一个元素的集合 (于是可以取这个唯一的元素为 $h(y)$). 构造一个这样的集合的可能性 (如果利用 Hilbert 运算这是显然的) 是熟知的**选择公理**. 至今还有数学家怀疑这个公理, 但是首先人们可以证明它在逻辑上跟其他公理是相容的 (K.Gödel, 1939), 其次它在逻辑上是独立的 (P. Cohen, 1963). 自然这些证明仅当针对一个适当的数学系统时才有意义.

注 9 定理 4 保证了 h 的存在性, 一般它不是唯一的. 例如取 $\mathrm{X}=\mathbf{R}, \mathrm{Y}=\mathbf{R}_+$, 而 $f(x)=x^2$, 这时 f 显然是满射的 (所有正实数有一个平方根). 除了其他的可能性, 我们有下列函数 h:

$$h_1(y)=\sqrt{y} \quad \text{对于 } y \geqslant 0,$$
$$h_2(y)=-\sqrt{y} \quad \text{对于 } y \geqslant 0,$$

$$h_3(y)=\begin{cases}+\sqrt{y}, & \text{如果 } y \geqslant 0 \text{ 是有理数},\\ -\sqrt{y}, & \text{如果 } y \geqslant 0 \text{ 是无理数}.\end{cases}$$

初学者可能会认为函数 h_3 (它实际上不可能用图像表示 —— 它的图在“曲线”这个词的朴素意义下不是一条“曲线”) 是稀奇古怪的, 不过从纯集合论的观点看, 它不比前两个差.

定理 5 设 $f: \mathrm{X} \to \mathrm{Y}$ 是一个映射. 以下条件是等价的:

a) f 是双射的;

b) 存在映射 $g, h: \mathrm{Y} \to \mathrm{X}$ 使得

$$g \circ f = j_{\mathrm{X}}, \quad f \circ h = j_{\mathrm{Y}}.$$

此外, 如果这些条件满足, 映射 g 和 h 是唯一的并且相等.

性质 a) 和 b) 的等价性直接从定理 3 和 4 推出, 因为“双射的”意即“单射的且满射的”. 为了证明存在仅一个函数 g, 仅一个函数 h, 只需证明所有的函数 g 等于所有的函数 h. 而根据定理 2 我们有

$$(g \circ f) \circ h = g \circ (f \circ h),$$

这被改写为 $j_{\mathrm{X}} \circ h = h = g \circ j_{\mathrm{Y}}$, 显然有

$$j_{\mathrm{X}} \circ h = h, \quad g \circ j_{\mathrm{Y}} = g,$$

这就完成了证明.

设

$$f: \mathrm{X} \to \mathrm{Y}$$

是一个双射, 则存在唯一的一个映射 $g: \mathrm{Y} \to \mathrm{X}$ 满足

$$g \circ f = j_{\mathrm{X}}, \quad f \circ g = j_{\mathrm{Y}},$$

即

$$g(f(x)) = x, \quad f(g(y)) = y.$$

我们称 g 是 f 的**逆映射**, 通常记作

$$\overset{-1}{f},$$

也经常 (不妥地) 记作 f^{-1}. 显然对于所有 $x \in \mathrm{X}$ 和所有 $y \in \mathrm{Y}$, 关系

$$y = f(x), \quad x = f^{-1}(y)$$

是等价的, 而如果 $\mathrm{G} \subset \mathrm{X} \times \mathrm{Y}$ 是 f 的图像, 那么 $\overset{-1}{f}$ 的图像是满足 $(x, y) \in \mathrm{G}$ 的序偶 (y, x) 的集合 (在 $\mathrm{X} = \mathrm{Y} = \mathbf{R}$ 这种经典情形, 这表明 $\overset{-1}{f}$ 的图像从 f 的图像通过关于第一分角线的对称而得到).



注 10 切记不要把"逆映射"中的"逆"理解为倒数. 常有人认为函数 x 的"逆函数"是 $1/x$, 这是不对的, 函数 x 的逆函数是它自己 (一般表述是"恒等映射"等于它的逆映射). 所以必须注意 $f^{-1}(x)$ 和 $(f(x))^{-1} = \dfrac{1}{f(x)}$ 的区别: 即使后者有定义, 两者一般也是不同的.

定理 6 设 $f : \mathrm{X} \to \mathrm{Y}$ 和 $g : \mathrm{Y} \to \mathrm{Z}$ 是两个映射, 如果 f 和 g 是单射的 (满射的), 则 $g \circ f$ 也是单射的 (满射的); 如果 f 和 g 是双射的, 则 $g \circ f$ 也是双射的, 并且有

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

最后, 如果 f 是双射的, 则 f^{-1} 也是双射的, 并且有

$$(f^{-1})^{-1} = f.$$

如果 f 和 g 是单射的, 因为 g 是单射的, 关系 $g(f(x')) = g(f(x''))$ 蕴含 $f(x') = f(x'')$, 又因为 f 是单射的, 故 $x' = x''$. 这就证明了 $g \circ f$ 是单射的.

如果 f 和 g 是满射的, 对于所有 $z \in \mathrm{Z}$, 存在 $y \in \mathrm{Y}$ 使得 $z = g(y)$, 而后存在 $x \in \mathrm{X}$ 使得 $y = f(x)$, 由此得 $z = g(f(x))$, 这就证明了 $g \circ f$ 是满射的.

因而如果 f 和 g 是双射的, 则 $g \circ f$ 也是双射的, 并且我们有

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ j_{\mathrm{Y}} \circ f = f^{-1} \circ f = j_{\mathrm{X}},$$

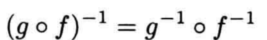
这证明 $g \circ f$ 的逆映射必然为 $f^{-1} \circ g^{-1}$.

最后, 如果 f 是双射的, 公式

$$f^{-1} \circ f = j_{\mathrm{X}}, \quad f \circ f^{-1} = j_{\mathrm{Y}}$$

表明 f^{-1} 同时是满射的和单射的 (对于 f^{-1} 援引定理 4 和 5), 并且以 f 作为其逆映射, 这就完成了证明.

注 11 "显然的"公式



$$(g \circ f)^{-1} = g^{-1} \circ f^{-1}$$

不但是错误的, 而且此式右端没有意义, 除非假定 $\mathrm{Y} = \mathrm{Z}$ (在这种情形下公式也是错误的).

9. 多变量函数

我们称所有的函数为**两个变量的函数**, 如果其出发集为两个集合的乘积或包含于一个这样的乘积. 如果 f 是一个两个变量的函数, 定义在乘积 $\mathrm{X} \times \mathrm{Y}$ 的一个子集 A 上, f 在一个点 (x, y) 的值本应记作 $f((x, y))$, 不过实际上用记号

$$f(x, y)$$

表示. 用类似方式定义三个, 四个, $\cdots$ 变量的函数, 并且采用记号 $f(x,y,z)$, $f(x,y,z,t)$, 等等.

数学上, 除了所使用的记号, 一个变量的函数和多个变量函数之间不存在任何差别 —— 由于事实上在第 3 小节关于函数的出发集没有做任何类别假设, 所有本节所叙述的内容可以无变化的应用到 “多个” 变量的函数. “一个” 和 “多个” 变量的差别来自于这样的事实, 到目前为止, 变量这个词表示的是今日所谓的 “实变量”, “一个变量的函数” 定义在实数集 $\mathbf{R}$ 的一个子集上, 至于三个变量的函数, 比如说, 它的出发集是 $\mathbf{R}^3$ 的一个子集. 我们在第 3 小节里给出一般定义, 它囊括了所有的函数的概念而没有任何例外, 正是为了避免这些区别.

在实际中往往要考虑出发集和到达集都是集合乘积的函数. 比如考虑一个映射

$$f : \mathrm{X} \times \mathrm{Y} \to \mathrm{U} \times \mathrm{V} \times \mathrm{W},$$

其中 X, Y, U, V, W 是任意集合. 考虑投影

$$\begin{aligned}\mathrm{pr}_1 &: \mathrm{U} \times \mathrm{V} \times \mathrm{W} \to \mathrm{U},\\ \mathrm{pr}_2 &: \mathrm{U} \times \mathrm{V} \times \mathrm{W} \to \mathrm{V},\\ \mathrm{pr}_3 &: \mathrm{U} \times \mathrm{V} \times \mathrm{W} \to \mathrm{W}.\end{aligned}$$

显然, 由投影的定义我们有

$$\text{对于所有 } z \in \mathrm{U} \times \mathrm{V} \times \mathrm{W},\ z = (\mathrm{pr}_1(z), \mathrm{pr}_2(z), \mathrm{pr}_3(z)).$$

考虑映射

$$\begin{aligned}f_1 &= \mathrm{pr}_1 \circ f : \mathrm{X} \times \mathrm{Y} \to \mathrm{U},\\ f_2 &= \mathrm{pr}_2 \circ f : \mathrm{X} \times \mathrm{Y} \to \mathrm{V},\\ f_3 &= \mathrm{pr}_3 \circ f : \mathrm{X} \times \mathrm{Y} \to \mathrm{W}.\end{aligned}$$

显然对于任意 $x \in \mathrm{X}$ 和 $y \in \mathrm{Y}$ 有

$$f(x,y) = (f_1(x,y), f_2(x,y), f_3(x,y)).$$

如此看来, 为了构造 f, 必须且只需知道定义在 $\mathrm{X} \times \mathrm{Y}$ 分别在 U, V, W 取值的三个函数. 在实际中我们记

$$f = (f_1, f_2, f_3)$$

(这个记号跟表示三元有序组的记号混淆, 但 “无大碍”).

例 5　设 X 和 Y 是两个集合, 考虑由

$$f(x,y) = (y,x)$$

给定的映射

$$f : \mathrm{X} \times \mathrm{Y} \to \mathrm{Y} \times \mathrm{X},$$

这显然是一个双射, 其逆映射就是 $(x,y)\to(y,x)$. 我们称 f 是**从 $\mathrm{X}\times\mathrm{Y}$ 到 $\mathrm{Y}\times\mathrm{X}$ 上的典范双射**. 当 $\mathrm{X}=\mathrm{Y}=\mathbf{R}$, f 就是 "关于第一分角线的对称".

例 6 设 X, Y, Z 是三个集合, 由

$$f((x,(y,z)))=((x,y),z)$$

给定的映射

$$f:\mathrm{X}\times(\mathrm{Y}\times\mathrm{Z})\to(\mathrm{X}\times\mathrm{Y})\times\mathrm{Z}$$

是双射. 我们这里也说这是**从 $\mathrm{X}\times(\mathrm{Y}\times\mathrm{Z})$ 到 $(\mathrm{X}\times\mathrm{Y})\times\mathrm{Z}$ 上的典范双射**. 正如在第 2 小节内说过的, 我们在实际中不区分 $(x,(y,z))$ 和 $((x,y),z)$.

§2 习题

1. 设 I 是满足条件 $0\leqslant\theta\leqslant 2\pi$ 的实数 θ 的集合, 而 G 是绕给定点 O 的旋转的集合. 考虑从 I 到 G 内的映射 f: 令每个数 $\theta\in\mathrm{I}$, 对应绕点 O 的角为 θ 的旋转. 映射 f 是满射吗? 是单射吗? 是双射吗? 当取 I 为满足条件 $0<\theta\leqslant 2\pi$ 的实数的集合时情况如何?

2. 设 X 和 Y 是集合, 要使 $\mathrm{X}\times\mathrm{Y}$ 的一个子集 G 是从 X 到 Y 内的一个映射的图, 必须并且只需从 G 到 X 内的映射 pr_1 是双射.

3. 设 $f:\mathrm{X}\to\mathrm{Y}$ 和 $g:\mathrm{Y}\to\mathrm{Z}$ 是两个映射, 而 $h=g\circ f$ 是复合映射.

a) 如果 h 是单射的, 则 f 是单射的; 如果 f 还是满射的, 则 g 是单射的.

b) 如果 h 是满射的, 则 g 是满射的; 如果 g 还是单射的, 则 f 是满射的.

4. 考虑集合 X, Y, Z 和映射

$$f:\mathrm{X}\to\mathrm{Y},\quad g:\mathrm{Y}\to\mathrm{Z},\quad h:\mathrm{Z}\to\mathrm{X},$$

组成复合映射

$$h\circ g\circ f,\quad g\circ f\circ h,\quad f\circ h\circ g,$$

假定或者它们中的两个是满射的, 而第三个是单射的, 或者它们中的两个是单射的, 而第三个是满射的. 证明 f,g 和 h 都是双射的.

5. 设 X, Y, Z 是三个集合, E 是从 $\mathrm{X}\times\mathrm{Y}$ 到 Z 内的所有映射的集合, 而 F 是从 X 到集合 Z^{Y} 的所有映射的集合, 这里 Z^{Y} 是从 Y 到 Z 内的所有映射的集合. 构造从 E 到 F 上的双射.

¶6. 称所有三元组

$$f=(\mathrm{G},\mathrm{X},\mathrm{Y}),\quad 其中\ \mathrm{G}\subset\mathrm{X}\times\mathrm{Y}$$

为 **X 和 Y 之间的对应**. 这个概念推广了从 X 到 Y 内的映射的概念 (X 和 Y 之间的对应, 也说是从 X 到 Y 的对应, 经常称为 "非处处定义的多值函数", 下面会说明这样称呼的理由. 最近才使用的这个术语, 显得不太妥当, 让人猜想对应的概念是函数概念的特殊情形, 其实反过来才是真的.) 集合 G 称为 f 的**图**. 说 **f 定义在 X 的 x 上**, 如果 $x\in\mathrm{pr}_1(\mathrm{G})$; 那么至少存在一个 $y\in\mathrm{Y}$, 使得 $(x,y)\in\mathrm{G}$, 我们说 **x 和 y 在 f 下相对应** (自然可能会遇到 f 不是对于所有的 $x\in\mathrm{X}$ 有定义, 如果 f 在 x 有定义, 也可能存在多个 $y\in\mathrm{Y}$, 它在 f 下对应到 x; 这两种情形解释了术语 "非处处定义的多值函数").

a) 对于实数集 $\mathrm{X} = \mathrm{Y} = \mathbf{R}$, 研究对应: 其图是由下列方程定义的集合:

$$xy = 1; \quad axy + bx + cy + d = 0; \quad x^2 + y^2 = 1; \quad x = \sin y$$

(在第二个方程中, a, b, c, d 是给定的实常数). 在每一个情形, 确定对应有定义的 x 的值, 以及对应于 x 的 y.

b) 设 Γ 和 Γ' 是两个不同的圆周; 取 Γ 的点的集合为 X 和 Y, 而 G 是这样的序偶 $(x, y) \in \mathrm{X} \times \mathrm{X}$, x, y 位于 Γ' 的同一条切线上. f 在 Γ 的哪些点有定义? 多少个点对应 Γ 的一个 f 有定义的点?

¶7. 设 $f = (\mathrm{G}, \mathrm{X}, \mathrm{Y})$ 是两个集合 X 和 Y 之间的一个对应 (参见习题 6). 对于 X 的所有子集 A, 用 $f(\mathrm{A})$ 记在 f 下对应到至少一个 $x \in \mathrm{A}$ 的 $y \in \mathrm{Y}$ 的集合, 而对于 Y 的所有子集 B, 用 $f^{-1}(\mathrm{B})$ 表示至少一个 $y \in \mathrm{B}$ 在 f 下对应到 x 的 $x \in \mathrm{X}$ 的集合. 证明, 如果 $\mathrm{A} \subset \mathrm{pr}_1(\mathrm{G})$, 则

$$\mathrm{A} \subset f^{-1}(f(\mathrm{A})).$$

如果条件 $\mathrm{A} \subset \mathrm{pr}_1(\mathrm{G})$ 不满足, 结论还成立吗? 又问, 在什么条件下, 关系

$$\mathrm{B} \supset f(f^{-1}(\mathrm{B}))$$

成立?

¶8. 设 $f = (\mathrm{G}, \mathrm{X}, \mathrm{Y})$ 和 $g = (\mathrm{H}, \mathrm{Y}, \mathrm{Z})$ 是两个对应. 称对应

$$g \circ f = (\mathrm{K}, \mathrm{X}, \mathrm{Z}) = h$$

为 g 和 f 的**复合**, 其定义是: $(x, z) \in \mathrm{K}$ 当且仅当存在 $y \in \mathrm{Y}$, 使得 $(x, y) \in \mathrm{G}$ 并且 $(y, z) \in \mathrm{H}$. 证明这个定义推广了两个映射的复合; 推广 §2 的定理 2 到对应. 给定 X 的一个子集 A, 必然有关系 $h(\mathrm{A}) = g(f(\mathrm{A}))$ 吗?

如果 $f = (\mathrm{G}, \mathrm{X}, \mathrm{Y})$ 是一个对应, 对应

$$f^{-1} = (\mathrm{G}', \mathrm{Y}, \mathrm{X})$$

称为 f 的**逆对应**, 其中 $\mathrm{G}' \subset \mathrm{Y} \times \mathrm{X}$ 是使得 $(x, y) \in \mathrm{G}$ 的序偶 (y, x) 的集合. 复合对应 $f^{-1} \circ f$ 是从 X 到 X 内的恒等映射吗? 证明, 如果 f 是从 X 到 Y 内的一个映射, 为了其逆对应 f^{-1} 本身是从 Y 到 X 内的一个映射, 必须并且只需 f 是双射的; 这时 f^{-1} 是 f 的逆映射.

设 $f = (\mathrm{G}, \mathrm{X}, \mathrm{Y})$ 和 $g = (\mathrm{H}, \mathrm{Y}, \mathrm{Z})$ 是两个对应. 公式

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$$

正确吗?

9. 设 f 是从一个集合 X 到一个集合 Y 内的映射. 假定

$$\text{关系 } f(x') \neq f(x'') \text{ 蕴含 } x' \neq x'',$$

证明 f 是单射.

§3 并集和交集

1. 两个集合的并集和交集

设 X 和 Y 是两个集合, 称记作

$$X \cap Y$$

的集合为 **X 和 Y 的交集**, 其定义是: 关系 $z \in X \cap Y$ 等价于关系

$$z \in X \text{ 且 } z \in Y.$$

换句话说, $X \cap Y$ 由同时属于 X 和 Y 的对象组成. 另外, 称记作

$$X \cup Y$$

的集合为 **X 和 Y 的并集**, 其定义是: 关系 $z \in X \cup Y$ 等价于关系

$$z \in X \text{ 或 } z \in Y,$$

换句话说, $X \cup Y$ 由或属于 X, 或属于 Y, 或同时属于 X 和 Y 的对象组成.

¶注 1 具有所指出的性质的集合 $X \cap Y$ 和 $X \cup Y$ 的存在性直观上是显然的, 但数学上看根本不是这样. $X \cap Y$ 的存在性借助于 §1 的定理 4 (把它用到 X 和关系 $x \in Y$). $X \cup Y$ 的存在性同样得到, 如果预先知道存在一个同时包含 X 和 Y 的集合 (对于这个集合, 以及关系 $z \in X$ 或 $z \in Y$, 应用 §1 的定理 4); 但是同时包含 X 和 Y 的集合的存在性是一个公理 (或是从用来构造一族集合的并集的更一般的公理得到的一个结果, 见第 2 小节), 于是努力从数学上证明上述集合的存在性是徒劳的.



显然有关系

$$X \cap Y \subset X, \quad Y \subset X \cup Y.$$

此外, 设 Z 是任意一个集合, Z 包含于 X 并且包含于 Y, 必须且只需对于所有的 $z \in Z$ 有 $z \in X$ 且 $z \in Y$, 即 $z \in X \cap Y$, 这也就是 $Z \subset X \cap Y$, 于是 $X \cap Y$ 是同时含于 X 和 Y 的最大的集合. 同样, Z 包含 X 并且包含 Y, 必须且只需 $Z \supset X \cup Y$, 于是 $X \cup Y$ 是同时包含 X 和 Y 的最小的集合.

对于两个集合 X, Y, 如果

$$X \cap Y = \varnothing,$$

即 X 和 Y 没有公共元素, 则称 **X 和 Y 是不交的**.

使用符号 $\cup$ 和 $\cap$ 计算的规则是十分简单的, 我们经常使用它们而不注明. 建议读者自己证明这些规则, 这里列举最主要的一些:

$$\mathrm{X} \cap \mathrm{Y} = \mathrm{Y} \cap \mathrm{X}, \quad \mathrm{X} \cup \mathrm{Y} = \mathrm{Y} \cup \mathrm{X},$$
$$\mathrm{X} \cap (\mathrm{Y} \cap \mathrm{Z}) = (\mathrm{X} \cap \mathrm{Y}) \cap \mathrm{Z}, \quad \mathrm{X} \cup (\mathrm{Y} \cup \mathrm{Z}) = (\mathrm{X} \cup \mathrm{Y}) \cup \mathrm{Z},$$
$$\mathrm{X} \cap (\mathrm{Y} \cup \mathrm{Z}) = (\mathrm{X} \cap \mathrm{Y}) \cup (\mathrm{X} \cap \mathrm{Z}),$$
$$\mathrm{X} \cup (\mathrm{Y} \cap \mathrm{Z}) = (\mathrm{X} \cup \mathrm{Y}) \cap (\mathrm{X} \cup \mathrm{Z}),$$
$$(\mathrm{X} - \mathrm{A}) \cap (\mathrm{X} - \mathrm{B}) = \mathrm{X} - (\mathrm{A} \cup \mathrm{B}), \quad \text{如果 } \mathrm{A}, \mathrm{B} \subset \mathrm{X}.$$

2. 一族集合的并集(*)

设 $(\mathrm{A}_i)_{i \in \mathrm{I}}$ 是一个集族 (§2, 第 3 小节, 注 4), 称如下定义的集合 A 为这个集族的并集: 关系 $x \in \mathrm{A}$ 等价于

$$\text{存在一个 } i \in \mathrm{I} \text{ 使得 } x \in \mathrm{A}_i.$$

当 I 仅含有两个元素, 比如记为 i 和 j 时, 显然并集就是前一小节定义的集合 $\mathrm{A}_i \cup \mathrm{A}_j$. 在任意指标集 I 的情形下, 并集的*存在性*是一个数学公理, 我们限于承认它. 反之, 并集的*唯一性* (即至多存在一个具有所指出的性质的集合 A 这一事实) 可以借助 §1 的定理 2 证明.



注 2 当谈到一个集族 $(\mathrm{A}_i)_{i \in \mathrm{I}}$ 时, 并未假定诸 A_i 是不依赖指标 i 的同一个集合的子集, 但并集的存在性指出事实上这个集合必然是存在的 (甚至更有: 考虑到 §1 的定理 4, 一个包含所有集 A_i 的集合的存在性等价于它们的并集的存在性).

为了表示一个集族 $(\mathrm{A}_i)_{i \in \mathrm{I}}$ 的并集, 使用记号

$$\bigcup_{i \in \mathrm{I}} \mathrm{A}_i.$$

虽然这个记号里涉及一个字母 i (直观地它表示 I 的一个 "变动元素"), 但结果中显然不依赖于 i, 并且可以在前面的记号里用任何其他未使用过的字母代替字母 i.

定理 1 *设 A 是一个集族 $(\mathrm{A}_i)_{i \in \mathrm{I}}$ 的并集. 一个集合 X 对于任意指标 $i \in \mathrm{I}$ 包含 A_i, 必须且只需 X 包含 A.*

假定 X 包含所有 A_i, 如果 $x \in \mathrm{A}$, 那么存在一个 i 使得 $x \in \mathrm{A}_i$, 由于 $\mathrm{A}_i \subset \mathrm{X}$, 故 $x \in \mathrm{X}$, 从而 $\mathrm{A} \subset \mathrm{X}$. 反之, 如果 X 包含 A, 为了证明 X 包含所有 A_i, 只需指出对于所有 i 有 $\mathrm{A} \supset \mathrm{A}_i$, 而这是显然的.

(*) 初次阅读时可以不读本小节和以下的小节, 或把它们当作习题.

定理 2 (并集的结合性) 设 $(\mathrm{A}_i)_{i\in\mathrm{I}}$ 和 $(\mathrm{I}_\lambda)_{\lambda\in\Lambda}$ 是集族, 并且假定

$$\mathrm{I} = \bigcup_{\lambda\in\Lambda} \mathrm{I}_\lambda,$$

则有

$$\bigcup_{i\in\mathrm{I}} \mathrm{A}_i = \bigcup_{\lambda\in\Lambda}\left(\bigcup_{i\in\mathrm{I}_\lambda} \mathrm{A}_i\right).$$

事实上, 令

$$\mathrm{B}_\lambda = \bigcup_{i\in\mathrm{I}_\lambda} \mathrm{A}_i,$$

x 属于集族 $(\mathrm{A}_i)_{i\in\mathrm{I}}$ 的并集, 必须且只需存在一个 $i\in\mathrm{I}$ 使得 $x\in\mathrm{A}_i$. 由于 I 是 I_λ 的并集, 这表明存在 $\lambda\in\Lambda$ 使得 $i\in\mathrm{I}_\lambda$, 于是存在 $\lambda\in\Lambda$ 使得 $x\in\mathrm{B}_\lambda$. 随之集族 $(\mathrm{A}_i)_{i\in\mathrm{I}}$ 的并集等于集族 $(\mathrm{B}_\lambda)_{\lambda\in\Lambda}$ 的并集, 这就完成了证明.

注 3 定理 2 说明, 计算一个并集, 可以把它的项分成若干组, 再把每个组用其并集取代.



定理 3 设 $f:\mathrm{X}\to\mathrm{Y}$ 是一个映射, 而 $(\mathrm{A}_i)_{i\in\mathrm{I}}$ 是 X 的子集的一个集族, 则有

$$f\left(\bigcup_{i\in\mathrm{I}}\mathrm{A}_i\right) = \bigcup_{i\in\mathrm{I}} f(\mathrm{A}_i).$$

对于 $y\in\mathrm{Y}$, 关系

$$y \in \bigcup_{i\in\mathrm{I}} f(\mathrm{A}_i)$$

等价于存在一个 $i\in\mathrm{I}$ 使得 $y\in f(\mathrm{A}_i)$, 即存在一个 $i\in\mathrm{I}$ 和一个 $x\in\mathrm{A}_i$ 使得 $y=f(x)$, 即存在一个 x 满足

$$y=f(x) \quad 和 \quad x\in\bigcup_{i\in\mathrm{I}}\mathrm{A}_i,$$

这就完成了证明.

3. 一族集合的交集

称如下定义的集合 A 为集合 $(\mathrm{A}_i)_{i\in\mathrm{I}}$ 的非空(*) 族的**交集**: 关系 $x\in\mathrm{A}$ 等价于关系

$$对于所有\ i\in\mathrm{I}\ 有\ x\in\mathrm{A}_i.$$

这个交集表示为

$$\bigcap_{i\in\mathrm{I}}\mathrm{A}_i.$$

(*) 这个条件意指 I 是非空的. 如果 I 是空集, 关系 “对于所有 $i\in\mathrm{I}$ 有 $x\in\mathrm{A}_i$” 对于任意 x 都是满足的, 从而不能定义一个集合 (否则就有了所有集合的集合).

定理 4 设 A 是集合的一个非空族 $(\mathrm{A}_i)_{i\in\mathrm{I}}$ 的交集. 一个集合 X 包含于 A, 必须且只需对于任意指标 $i\in\mathrm{I}$, X 包含于 A_i.

证明类似于定理 1 的证明, 留给读者作为习题.

定理 5 (交集的结合性) 设 $(\mathrm{A}_i)_{i\in\mathrm{I}}$ 和 $(\mathrm{I}_\lambda)_{\lambda\in\Lambda}$ 是两个集族, 假定 I,Λ 和诸 I_λ 非空, 并且假定

$$\mathrm{I}=\bigcup_{\lambda\in\Lambda}\mathrm{I}_\lambda;$$

则有

$$\bigcap_{i\in\mathrm{I}}\mathrm{A}_i=\bigcap_{\lambda\in\Lambda}\left(\bigcap_{i\in\mathrm{I}_\lambda}\mathrm{A}_i\right).$$

证明类似于定理 2 的证明.

定理 6 设 $f:\mathrm{X}\to\mathrm{Y}$ 是一个映射, 而 $(\mathrm{A}_i)_{i\in\mathrm{I}}$ 是 X 的子集的一个非空集族, 则

$$f\left(\bigcap_{i\in\mathrm{I}}\mathrm{A}_i\right)\subset\bigcap_{i\in\mathrm{I}}f(\mathrm{A}_i);$$

如果 f 是单射的, 则

$$f\left(\bigcap_{i\in\mathrm{I}}\mathrm{A}_i\right)=\bigcap_{i\in\mathrm{I}}f(\mathrm{A}_i).$$

设对于所有的 i 有 $x\in\mathrm{A}_i$, 则有 $f(x)\in f(\mathrm{A}_i)$, 这就证明了定理的第一个结论. 现在假定 f 是单射的, 考虑 $f(\mathrm{A}_i)$ 的交集的一个元素 y. 对于所有的 i 存在 A_i 的一个元素 x_i 使得 $y=f(x_i)$, 但由于 f 是单射的, 仅存在唯一的 x 使得 $y=f(x)$, 于是对于所有的 i 必然有 $x=x_i$; 从而对于所有的 i 有 $x\in\mathrm{A}_i$, y 属于 A_i 的交集在 f 下的像, 于是有

$$\bigcap_{i\in\mathrm{I}}f(\mathrm{A}_i)\subset f\left(\bigcap_{i\in\mathrm{I}}\mathrm{A}_i\right),$$

这就证明了定理的第二个结论, 因为在任何情形下已经有了反向的包含关系.

注 4 如果 f 不是单射, 则上述定理的第二个结论可能是错误的. 作为例子, 取至少含有两个点 a 和 b 的集合作为 Y, 取乘积 $\mathrm{Y}\times\mathrm{Y}$ 作为 X, 而取 pr_2 作为 f. 设 A 是序偶 (a,y) 的集合, 其中 $y\in\mathrm{Y}$, 而 B 是序偶 (b,y) 的集合, 其中 $y\in\mathrm{Y}$. 显然有 $\mathrm{A}\cap\mathrm{B}=\varnothing$, 故 $f(\mathrm{A}\cap\mathrm{B})=\varnothing$; 但是 $f(\mathrm{A})=f(\mathrm{B})=\mathrm{Y}$, 故 $f(\mathrm{A})\cap f(\mathrm{B})$ 非空.

定理 7 设 $f:\mathrm{X}\to\mathrm{Y}$ 是一个映射, 而 $(\mathrm{A}_i)_{i\in\mathrm{I}}$ 是 Y 的子集的非空族. 则有

$$\overset{-1}{f}\left(\bigcap_{i\in\mathrm{I}}\mathrm{A}_i\right)=\bigcap_{i\in\mathrm{I}}\overset{-1}{f}(\mathrm{A}_i).$$

事实上, $x \in X$ 属于左端, 必须并且只需 $f(x)$ 属于 A_i 的交集, 即对于所有的 i 有 $f(x) \in A_i$, 换句话说, 对于所有的 i 有 $x \in f^{-1}(A_i)$, 或有 $x \in X$ 属于右端, 由此得到定理.

同样证明公式

$$\overset{-1}{f}\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} \overset{-1}{f}(A_i).$$

定理 8 设 $(A_i)_{i\in I}$ 是集合 X 的子集的非空族. 则有

$$X - \bigcap_{i\in I} A_i = \bigcup_{i\in I}(X - A_i);\quad X - \bigcup_{i\in I} A_i = \bigcap_{i\in I}(X - A_i).$$

事实上, 设 $x \in X$, 关系

$$x \in X - \bigcap_{i\in I} A_i$$

等价于关系

对于所有 $i \in I$ 有 $x \in A_i$

的否定, 故等价于关系

存在一个 $i \in I$, 使得 $x \notin A_i$,

随之等价于

存在一个 $i \in I$, 使得 $x \in X - A_i$,

故等价于

$$x \in \bigcup_{i\in I}(X - A_i),$$

这就证明了第一个公式. 考虑到对于 X 的所有子集皆成立的关系 $X - (X - A) = A$, 第二个公式可以由第一个公式推导出来.

§3 习题

1. 设 $(A_i)_{i\in I}$ 是 X 的子集的一个族, 而 $(B_j)_{j\in J}$ 是 Y 的子集的一个族. 借助这些族的交集和并集计算 $X \times Y$ 的子集的族 $(A_i \times B_j)_{i\in I, j\in J}$ 的交集和并集.

2. 设 $(A_i)_{i\in I}$ 是 X 的子集的一个族, 于是对于每一个 $i \in I$ 有 $\mathscr{P}(A_i) \subset \mathscr{P}(X)$. 公式

$$\mathscr{P}\left(\bigcap_{i\in I} A_i\right) = \bigcap_{i\in I} \mathscr{P}(A_i),$$

$$\mathscr{P}\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} \mathscr{P}(A_i)$$

成立吗?

3. 设 X 和 Y 是两个集合, 而 $(X_i)_{i\in I}$ 是 X 的子集的一个族, 其并集是整个集合 X. 假定对于每个 i 给定一个从 X_i 到 Y 内的映射 f_i. 指出以下两个条件是等价的:

a) 对于任意 $i, j \in \mathrm{I}$, $f_i(x) = f_j(x)$ 对于所有 $x \in \mathrm{X}_i \cap \mathrm{X}_j$ 成立;

b) 存在从 X 到 Y 内的一个映射 f, 对于所有 $i \in \mathrm{I}$, f 在 X_i 上与 f_i 重合 (人们说 f 由 f_i **黏合**而得). 这样的 f 是唯一的.

4. 称 X 的子集的一个族 $(\mathrm{U}_i)_{i\in\mathrm{I}}$ 是 X 的一个**覆盖**, 如果这个族的并集是这个 X. 给定 X 的两个覆盖 $(\mathrm{U}_i)_{i\in\mathrm{I}}$ 和 $(\mathrm{V}_j)_{j\in\mathrm{J}}$, 如果对于所有的指标 $j \in \mathrm{J}$, 存在一个指标 $i \in \mathrm{I}$, 使得 $\mathrm{V}_j \subset \mathrm{U}_i$, 则说第二个覆盖比第一个覆盖**更细**.

证明给定 X 的两个覆盖, 则存在 X 的比这两个覆盖更细的第三个覆盖.

¶5. 称一个集合 K 和 K 的非空有限子集 (称为所考虑的单纯图的**单形**) 的一个集合组成的一个对象为一个**单纯图**, 这些有限子集满足下列条件: K 的一个单形的所有非空子集仍然是 K 的一个单形. [一个单纯图不仅仅是一个集合 K, 这是一个由 K 和 K 的子集的一个集合组成的序偶; 不过总是用对应的集合同样的字母表示一个单纯图.] 在一个单纯图 K 里, 含有 $n+1$ 个元素的所有单形称为 n **维单形**; 0 维单形称为 K 的**顶点**, 1 维单形称为 K 的**边**, 2 维单形称为 K 的**面**, 等等. 单纯图的概念特别为研究任意维的多面体的拓扑性质而引入, 并且已经出现在 Euler 的著作中, 他证明了在通常空间内一个多面体的表面, 如果有 a 个顶点, b 个棱和 c 个面, 则数 $a-b+c$ 不依赖把这个面分成三角形的方式. 从一个多面体表面分成三角形出发, 我们可以构造一个单纯图如下: 集合 K 是所考虑的多面体 P 的顶点的集合; K 的所有一个元素的子集是 K 的一个顶点; K 的两个元素的子集 $\{a, b\}$ 是 K 的一条边, 当且仅当连接 a 到 b 的直线段是 P 的一个棱; 最后, K 的三个元素的集合 $\{a, b, c\}$ 是 K 的一个面, 当且仅当顶点为 a, b, c 的三角形是多面体 P 的一个面. 推广到任意维数是容易的.

a) 设 X 是任意一个集合, 而 $(\mathrm{U}_i)_{i\in\mathrm{I}}$ 是 X 的一个覆盖. 我们约定说指标集 I 的一个子集 S 是一个单形, 如果它是有限的, 非空的, 并且交集

$$\bigcap_{i\in\mathrm{S}} \mathrm{U}_i$$

是非空的. 指出集合 I 和刚定义的单形的集合组成的序偶是一个单纯图 (称为所考虑的覆盖的**神经**).

b) 设 K 是一个单纯图. 用 P(K) 表示定义在 K 上的取实数值的函数 f 的集合, f 具有下列三个性质: 使得 $f(x) \neq 0$ 的 $x \in \mathrm{K}$ 的集合是 K 的一个单形; 对于所有 $x \in \mathrm{K}$ 有 $f(x) \geqslant 0$; f 在 K 的不同的点的值的和

$$\sum_{x\in\mathrm{K}} f(x)$$

(这个和仅含有有限个非零项) 是 1. 最后, 对于所有 $x \in \mathrm{K}$, 用 U_x 表示使得 $f(x) \neq 0$ 的 $f \in \mathrm{P(K)}$ 的集合. 指出 $(\mathrm{U}_x)_{x\in\mathrm{K}}$ 是集合 P(K) 的一个覆盖, 并且这个覆盖的神经恰好是给定的单纯图 K.

[如果单纯图 K 是有限的, 并且由 n 个元素 $x_1, \cdots, x_n$ 组成, 可以用 $\mathbf{R}^n$ 的其坐标为 n 个数 $f(x_1), \cdots, f(x_n)$ 的点表示; 我们得到从 P(K) 到 $\mathbf{R}^n$ 的一个多面体上的双射, 多面体的 "形状" 恰好依赖在给定的单纯图的各个单形之间存在的组合关系. 这个基本的操作让我们把看起来纯粹 "定性" 的问题, 即多面体 "形状" 的研究, 变换为一个纯粹的代数问题, 即有限单纯图的研究.]

§4 等价关系

1. 等价关系

设 R 是涉及两个变量 xy 的关系. 我们说这是一个**等价关系**, 如果下列条件满足:

a) 对于所有 x, $\mathrm{R}\{x,x\}$ 是真的;

b) 关系 $\mathrm{R}\{x,y\}$ 蕴含关系 $\mathrm{R}\{y,x\}$;

c) $\mathrm{R}\{x,y\}$ 和关系 $\mathrm{R}\{y,z\}$ 蕴含关系 $\mathrm{R}\{x,z\}$.

显然 $x=y$ 是一个等价关系.

与此概念类似且在实际中更有用的是在一个**集合 E 上的等价关系**, 这是涉及两个变量 xy 的满足下列条件的关系:

(R0) 关系 $\mathrm{R}\{x,y\}$ 蕴含 $x\in\mathrm{E}$ 和 $y\in\mathrm{E}$;

(R1) 对于所有 $x\in\mathrm{E}$, 关系 $\mathrm{R}\{x,x\}$ 是真的;

(R2) 关系 $\mathrm{R}\{x,y\}$ 蕴含关系 $\mathrm{R}\{y,x\}$;

(R3) $\mathrm{R}\{x,y\}$ 和关系 $\mathrm{R}\{y,z\}$ 蕴含关系 $\mathrm{R}\{x,z\}$.

如果 R 是集合 E 上的等价关系, 称使得关系 $\mathrm{R}\{x,y\}$ 为真的序偶 $(x,y)\in\mathrm{E}\times\mathrm{E}$ 的集合 $\mathrm{G}\subset\mathrm{E}\times\mathrm{E}$ 为 **R 的图**. 于是关系 $\mathrm{R}\{x,y\}$ 是真的等价于 $(x,y)\in\mathrm{G}$, 条件 (R1) 表明 G 包含乘积 $\mathrm{E}\times\mathrm{E}$ 的对角集.

为了提供一个例子 (后面会指出它导出集合 E 上的**所有**等价关系), 考虑从 E 到一个任意集合 M 内的映射 f, 取关系

$$f(x)=f(y)$$

作为 $\mathrm{R}\{x,y\}$, 我们显然得到 E 上的一个等价关系, 称为**伴随于映射 f 的等价关系**.

现在给出另外一些例子.

例 1 对于所有的集合 E, $\mathrm{E}\times\mathrm{E}$ 的对角集 (§2, 第 7 小节, 例 2) 是 E 上的一个等价关系的图, 这个关系就是

$$x=y.$$

例 2 对于所有的集合 E, $\mathrm{E}\times\mathrm{E}$ 是 E 上的一个等价关系的图, 这个关系就是 $(x,y)\in\mathrm{E}\times\mathrm{E}$; 在这种情形, 关系 $\mathrm{R}\{x,y\}$ 对于所有 $x\in\mathrm{E}$ 和 $y\in\mathrm{E}$ 都是真的.

例 3 我们称集合 X **等势于**集合 Y, 如果存在从 X 到 Y 上的一个双射. 关系 "X 等势于集合 Y" 是一个等价关系; 事实上, X 等势于 X (考虑恒等映射 j_{X}, 这是双射); 如果 X 等势于集合 Y, 那么 Y 等势于集合 X (因为如果 f 是从 X 到 Y 上的一个双射, 那么 f^{-1} 是从 Y 到 X 上的一个双射); 最后, 如果 X 等势于 Y, Y 等势于 Z, 那么 X 等势于 Z (因为如果 f 是从 X 到 Y 上的一个双射, 而 g 是从 Y 到 Z 上的一个双射, 那么根据 §2 的定理 7, $g\circ f$ 是从 X 到 Z 上的一个双射). 参见下一节.

例 4 设 $\mathbf{Z}$ 是有理整数的集合, 即正的, 负的整数和 0 的集合,

$$\cdots,-2,-1,0,1,2,\cdots$$

选取一个整数 $p \geqslant 1$ 并且考虑 $\mathbf{Z}$ 的两个元素 x 和 y 之间的关系

$$p \text{ 整除 } x-y,$$

这是 $\mathbf{Z}$ 上的一个等价关系. 事实上, 关系 "p 整除 $x-x$" 总是真的; 关系 "p 整除 $x-y$" 显然蕴含关系 "p 整除 $y-x$"; 最后, 如果 "p 整除 $x-y$", 并且 "p 整除 $y-z$", 显然 p 整除 $x-z=(x-y)+(y-z)$.

这样得到的 $\mathbf{Z}$ 上的关系称为**模 p 同余**, 经典的记法是

$$x \equiv y \pmod{p},$$

读作 "x 同余于 y 模 p", 这表明 x 和 y 仅相差 p 的一个倍数. 两个世纪以来同余理论在数论里起着基本的作用.

例 5 写出

$$\text{存在一个整数 } n \text{ 使得 } x-y=2n\pi,$$

就在 $\mathbf{R}$ 上得到一个等价关系, 称为**模 2π 同余**. 这个等价关系是角的**数学**定义的基础.

例 6 取 E 是序偶 (p,q) 的集合, $p,q \in \mathbf{Z}$, $q \neq 0$, 对于 $x=(p,q), y=(p',q')$, 把关系

$$pq'=p'q$$

记为 $\mathrm{R}\{x,y\}$, 这样就得到 E 上的一个等价关系 (这不甚显然, 但是借助初等计算可以验证). 在后面 (§29) 将会看到, 这个等价关系是从整数出发的有理数定义的基础. 建议读者作为习题自行验证 R 是所考虑的 E 上的一个等价关系.

例 7 取 E 是通常空间的点的集合, 选定一条直线 D, 考虑关系

$$\text{存在过 } x \text{ 和 } y \text{ 的平行或重合于 D 的一条直线},$$

这样就得到 E 上的一个等价关系.

例 8 取 E 为平面上的三角形的集合, 取关系

$$\text{三角形 } x \text{ 和 } y \text{ 是全等的}$$

作为 $\mathrm{R}\{x,y\}$. 三角形的全等在初等几何里已经给了所谓定义, 即存在一个变换 x 成 y 的位移 (假定知道什么是位移), 这就得到 E 上的一个等价关系.

设 R 是集合 E 上的一个等价关系, 受例 4 记号的启发, 经常把 $R\{x,y\}$ 写成

$$x \equiv y \pmod{R}.$$

我们有下列性质: 关系

$$x \equiv x \pmod{R}$$

对于所有 $x \in E$ 成立; 关系

$$x \equiv y \pmod{R}$$

蕴含关系

$$y \equiv x \pmod{R};$$

最后有关系

$$x \equiv y \pmod{R} \quad 和 \quad y \equiv z \pmod{R}$$

蕴含关系

$$x \equiv z \pmod{R}.$$

2. 集合关于一个等价关系的商集

我们将要证明一个结果, 它指出如何得到一个集合上的所有可能的等价关系:

定理 1 设 R 是集合 E 上的一个等价关系, 则存在一个集合 M 和一个映射 $f: E \to M$ 使得关系

$$x \equiv y \pmod{R}$$

和

$$f(x) = f(y)$$

是等价的.

我们不仅要证明 f 和 M 的存在性, 而且要明确构造一个满足所述的条件的集合 M 和一个映射 f.

给定一个 $x \in E$, 称由使得

$$x \equiv y \pmod{R}$$

为真的 $y \in E$ 组成的集合 $F_x \subset E$ 为 x **模 R 的等价类**, 于是关系

$$x \equiv y \pmod{R}, \ y \in F_x$$

是等价的. 我们要证明对于 $x, y \in E$, 关系

$$x \equiv y \pmod{R}$$

和

$$\mathrm{F}_x = \mathrm{F}_y$$

是等价的.

事实上, 假定 $x \equiv y \pmod{\mathrm{R}}$, 那么关系 $z \in \mathrm{F}_y$ 意即 $y \equiv z \pmod{\mathrm{R}}$, 这蕴含 $x \equiv z \pmod{\mathrm{R}}$, 从而 $z \in \mathrm{F}_x$, 因此关系 $x \equiv y \pmod{\mathrm{R}}$ 蕴含 $\mathrm{F}_y \subset \mathrm{F}_x$. 同样有 $\mathrm{F}_x \subset \mathrm{F}_y$, 即得 $\mathrm{F}_x = \mathrm{F}_y$. 反之假定 $\mathrm{F}_x = \mathrm{F}_y$, 由于总有 $y \equiv y \pmod{\mathrm{R}}$, 从而 $y \in \mathrm{F}_y$, 由此得 $y \in \mathrm{F}_x$, 故 $x \equiv y \pmod{\mathrm{R}}$, 我们的断言被证.

现在我们可以证明定理 1. 在 E 的子集的集合 $\mathscr{P}(\mathrm{E})$ 里考虑 E 的子集 F 组成的集合, 这里要求 F 满足条件: 对于至少一个 $x \in \mathrm{E}$ 有

$$\mathrm{F} = \mathrm{F}_x,$$

这是 E 的不同元素的等价类的集合, 记为 E/R, 称为 **E 关于等价关系 R 的商集**. 现在由

$$f(x) = \mathrm{F}_x$$

定义一个映射

$$f : \mathrm{E} \to \mathrm{E/R},$$

即令每个 $x \in \mathrm{E}$ 对应它模 R 的等价类. 上面已经看到关系

$$x \equiv y \pmod{\mathrm{R}}$$

蕴含 $\mathrm{F}_x = \mathrm{F}_y$; 但是这意味着

$$f(x) = f(y),$$

定理得证.

刚才定义的映射 f 称为**从 E 到 E/R 上的典范映射**. 由 E/R 的构造本身知道 f 显然是满射的.

注意类 F_x 具有两个重要性质: F_x 的并集是整个 E (由于事实: 对于所有 $x \in \mathrm{E}$ 有 $x \in \mathrm{F}_x$); 另外, 任意两个类 F_x 和 F_y 或者相等或者是不交的, 因为如果 $\mathrm{F}_x \cap \mathrm{F}_y$ 至少含有一个元素 z, 那么就有关系

$$x \equiv z \pmod{\mathrm{R}} \quad 和 \quad y \equiv z \pmod{\mathrm{R}},$$

由此利用 (R2) 和 (R3) 即得关系

$$x \equiv y \pmod{\mathrm{R}},$$

在定理 1 的证明中已经看到这蕴含 $\mathrm{F}_x = \mathrm{F}_y$.

注 1 定理 1 的证明方法的基础在于构造一个从集合 E 到 $\mathrm{F}_x = \mathrm{F}_y$ 的子集的集合 $\mathscr{P}(\mathrm{E})$ 内的映射, 即 $x \to \mathrm{F}_x$. 如果我们仅了解古典函数 (一个实变量、取值为实数的函数), 显然我们不能够实现这类构造. 这类证明诠释了考虑其出发集和到达集是任意集合的函数的必要性.

还要注意, 可能让初学者感到惊奇的是, E/R 的构造方法还在日常生活中被使用, 像下面的 (非数学的) 例子所表明的: 取 E 为人的集合, 作为 R 是关系: "x 和 y 是同民族的", 我们显然得到 E 上的一个等价关系. 对于 $x \in \mathrm{E}$, F_x 是所有跟 x 同民族的人的集合, 或者是 x 所属的民族; 随之商集 E/R 是现存的各个民族的总体, 而从 E 到 E/R 上的典范映射就是让每个人对应他 (或她) 所属的民族.

现在给出几个构造商集的例子.

例 9 考虑例 4 中的 $\mathbf{Z}$ 上的 (模 p 同余) 等价关系. 对于所有 $x \in \mathbf{Z}$, 类 F_x 显然由形如 $x + np$ 的整数组成, 其中 n 是任意一个整数. 这些类刚好有 p 个 (假定 $p \geqslant 1$). 事实上, 对于所有 $x \in \mathbf{Z}$ 可以写出 ("x 除以 p 的带余除法")

$$x = np + r \quad (0 \leqslant r < p),$$

显然 x 被 p 除的余数确定类 F_x; 换言之, 模 p 的每个类是下列各类之一:

$$\mathrm{F}_0, \mathrm{F}_1, \cdots, \mathrm{F}_{p-1};$$

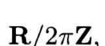
进而这 p 个类是两两不同的, 因为如果介于 0 和 $p-1$ 之间的整数 r 和 r' 关于 p 同余, 那么它们显然相等.

这里我们用

$$\mathbf{Z}/p\mathbf{Z}$$

表示 E/R, 它恰由 p 个元素组成, 这些元素称为**模 p 整数**. 一个模 p 整数其实就是普通整数的一个集合, 即由给定的一个整数 x 加上 p 的任意倍数所得到的所有整数的集合, 我们说整数 x 是所考虑的模 p 整数的一个**代表**. 所有模 p 整数有唯一的一个介于 0 和 $p-1$ 之间的代表, 这样就可以借助整数 $0, 1, \cdots, p-1$ 给模 p 整数编号, 并且代表它们. 比如, 用 0, 1, 2, 3, 4, 5 代表 $\mathbf{Z}/6\mathbf{Z}$ 的元素, 当我们说 $\mathbf{Z}/6\mathbf{Z}$ 的元素 4, 这就是意味人们在说 "对于模 6 同余整数 4 的类", 或 "形如 $6n+4$ 整数的集合". 利用这些语言的约定, 从 $\mathbf{Z}$ 到 $\mathbf{Z}/p\mathbf{Z}$ 上的典范映射就是令每个整数对应它被 p 除的余数.

例 10 在第 1 小节的例 5 里, 商集记作



$$\mathbf{R}/2\pi\mathbf{Z},$$

它的元素称为**模 2π 实数**. 一个模 2π 实数就是实数的一个集合, 即由一个给定的实数加上 2π 的任意整倍数得到的集合. 显然一个角的度量正是一个模 2π 的实数 (而非一个通常的实数).

例 11 在第 1 小节的例 7 里, 类是平行或重合于 D 的直线, 而商集是平行于 D 的直线的集合.

3. 定义在商集上的函数(*)

这里是一个十分有用的结果:

定理 2 设 E 是一个集合, 而 R 是 E 上的一个等价关系, p 是从 E 到 E/R 上的典范映射, f 是从 E 到一个集合 X 内的映射. 下列条件是等价的:

a) 关系 $x \equiv y \pmod{\mathrm{R}}$ 蕴含 $f(x) = f(y)$;

b) 存在一个映射

$$\overline{f} : \mathrm{E/R} \to \mathrm{X}$$

使得 $f = \overline{f} \circ p$.

如果这些条件满足, 映射 $\overline{f}$ 是唯一的. 它是单射的, 必须且只需关系 $x \equiv y \pmod{\mathrm{R}}$ 和 $f(x) = f(y)$ 是等价的; 它是满射的, 必须且只需 f 是满射的.

根据 §2 定理 1, 存在一个满足所述条件的映射 $\overline{f}$ 的必要和充分条件是关系 $p(x) = p(y)$ 蕴含关系 $f(x) = f(y)$; 而关系 $p(x) = p(y)$ 等价于关系 $x \equiv y \pmod{\mathrm{R}}$, 从而性质 a) 和 b) 是等价的.

$\overline{f}$ 的唯一性来自 p 是满射的. 事实上, 设找到两个映射 g 和 h 使得 $f = g \circ p = h \circ p$, 对于所有 $x \in \mathrm{E}$ 有 $g(p(x)) = h(p(x))$, 于是 g 和 h 在像 $p(\mathrm{E})$ 上重合, 即在 E/R 上重合, 即 $g = h$.

我们有 $f(\mathrm{E}) = \overline{f}(p(\mathrm{E})) = \overline{f}(\mathrm{E/R})$, 于是 f 是满射的, 当且仅当 $\overline{f}$ 是满射的. 为了映射 f 是单射的, 由于 E/R 的所有元素有形式 $p(z)$, 必须且只需

$$\text{关系 } \overline{f}(p(x)) = \overline{f}(p(y)) \text{ 蕴含 } p(x) = p(y),$$

换句话说 $f(x) = f(y)$ 蕴含关系 $x \equiv y \pmod{\mathrm{R}}$, 这就完成了证明.

下一个定理类似于前一个定理, 不过稍微复杂一些, 当要在一个商集上定义代数运算时, 它是基本的:

定理 3 设 X, Y, Z 是三个集合, R, S, T 是这些集合上的等价关系, 而 f 是从 X × Y 到 Z 内的一个映射. 用 $x \to \bar{x}, y \to \bar{y}, z \to \bar{z}$ 表示从 X, Y, Z 到 X/R, Y/S, Z/T 上的典范映射. 下列断言是等价的:

(*) 本小节的结果在 §29 之前不是必需的.

a) 关系 $x' \equiv x'' \pmod{R}$ 和 $y' \equiv y'' \pmod{S}$
蕴含关系

$$f(x', y') \equiv f(x'', y'') \pmod{T};$$

b) 存在一个映射

$$\overline{f} : (X/R) \times (Y/S) \to (Z/T)$$

使得对于任何 $x \in X$ 和 $y \in Y$ 有

$$\bar{f}(\bar{x}, \bar{y}) = \overline{f(x, y)}.$$

如果这些条件满足, 则映射 $\overline{f}$ 是唯一的.

考虑映射 $u : X \times Y \to Z/T$, 其定义是

$$u(x, y) = \overline{f(x, y)};$$

再考虑映射 $v : X \times Y \to (X/R) \times (Y/S)$, 其定义是

$$v(x, y) = (\bar{x}, \bar{y}).$$

所有的事情都归结为构造一个映射 $\overline{f} : (X/R) \times (Y/S) \to (Z/T)$, 使得 $u = \overline{f} \circ v$. 为此应用 §2 的定理 1: $\overline{f}$ 存在当且仅当

$$v(x', y') = v(x'', y'') \quad \text{蕴含} \quad u(x', y') = u(x'', y'');$$

而第一个关系表示为 $\overline{x'} = \overline{x''}$ 和 $\overline{y'} = \overline{y''}$, 即

$$x' \equiv x'' \pmod{R} \quad \text{和} \quad y' \equiv y'' \pmod{S},$$

第二个关系表示为

$$f(x', y') \equiv f(x'', y'') \pmod{T};$$

因此 §2 的定理 1 保证条件 a) 和 b) 的等价性.

如果存在一个映射 $\overline{f}$ 具有所要求的性质, 因为 v 显然是满射的, 故 $\overline{f}$ 是唯一的. 定理得证.

例 12 取 $X = Y = Z = \mathbf{Z}$ 为有理整数集合, 取 R, S 和 T 为模 p 的同余关系, 这里 p 是一个给定的正整数. 考虑由

$$f(x, y) = x + y, \quad g(x, y) = xy$$

给定的映射 $f, g : \mathbf{Z} \times \mathbf{Z} \to \mathbf{Z}$. 定理 3 的条件 a) 对于 f 和 g 是满足的, 为了确信这个断言, 我们应当验证关系

$$x' \equiv x'' \pmod{p} \quad \text{和} \quad y' \equiv y'' \pmod{p}$$

蕴含关系

$$x' + y' \equiv x'' + y'' \pmod{p} \quad 和 \quad x'y' \equiv x''y'' \pmod{p}.$$

根据等式

$$(x' + y') - (x'' + y'') = (x' - x'') + (y' - y''),$$
$$x'y' - x''y'' = (x' - x'')y' + x''(y' - y''),$$

这是显然的. 于是可以利用定理 3 到 f 和 g, 换句话说, 存在映射

$$\overline{f}, \overline{g} : (\mathbf{Z}/p\mathbf{Z}) \times (\mathbf{Z}/p\mathbf{Z}) \to (\mathbf{Z}/p\mathbf{Z}),$$

由下列性质刻画它们: 给定 $\mathbf{Z}/p\mathbf{Z}$ 的元素 ξ 和 η, 其代表是 x 和 y, 于是按照定理 3 的记号有

$$\xi = \bar{x}, \quad \eta = \bar{y},$$

那么 $\overline{f}(\xi, \eta)$ 和 $\overline{g}(\xi, \eta)$ 就有代表 $x + y$ 和 xy, 它们对于模 p 的类仅依赖 x 和 y 对于模 p 的类, 而不依赖所考虑的类 ξ 和 η 中的 x 和 y 的选取.

在实际中, 我们写成

$$\bar{f}(\xi, \eta) = \xi + \eta, \quad \bar{g}(\xi, \eta) = \xi\eta,$$

并且说 $\xi + \eta$ 和 $\xi\eta$ 分别是 $\mathbf{Z}/p\mathbf{Z}$ 的元素 ξ 和 η 的**和与乘积**. 如果用 θ 表示从 $\mathbf{Z}$ 到 $\mathbf{Z}/p\mathbf{Z}$ 的典范映射, 我们就会看到对于模 p 整数的加法和乘法的定义使得对于任意 $x, y \in \mathbf{Z}$ 有关系

$$\theta(x + y) = \theta(x) + \theta(y), \quad \theta(xy) = \theta(x)\theta(y).$$

实际上, 在 $\mathbf{Z}/p\mathbf{Z}$ 内的加法和乘法如下计算: 用普通整数 $0, 1, \cdots, p - 1$ 表示模 p 整数 (见例 9), 如果模 p 的类用 (介于 0 和 $p - 1$ 之间的)x 和 y 表示, 那么这些类的和与乘积将用 $x + y$ 和 xy 被 p 除所得的余数表示.

例如, 这里是模 5 整数的 “加法表和乘法表”:

	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

像 $2 \cdot 3 = 1$ 这样的等式意指

$$2 \cdot 3 \equiv 1 \pmod{5}.$$

除 $p = 1$ 外最简单的情形是 $p = 2$, 此时有两类, 按规定称为 “偶” 和 “奇”, 而计算规则由以下公式给出:

$$\begin{array}{ll} 偶 + 偶 = 偶 & 偶 \times 偶 = 偶 \\ 偶 + 奇 = 奇 & 偶 \times 奇 = 偶 \\ 奇 + 偶 = 奇 & 奇 \times 偶 = 偶 \\ 奇 + 奇 = 偶 & 奇 \times 奇 = 奇 \end{array}$$

§4 习题

1. 称所有其并集为集合 X 的、两两不交的非空集合的族 $(\mathrm{A}_i)_{i\in \mathrm{I}}$ 为集合 X 的一个**划分**. 给定一个这样的划分, 考虑 X 的元素 x, y 之间的关系

$$存在一个\ i \in \mathrm{I},\ 使得\ x \in \mathrm{A}_i,\ 并且\ y \in \mathrm{A}_i.$$

指出这个关系是一个等价关系, 并由这个关系构造类和商集. 指出 X 上的所有等价关系可以用这种方法得到.

¶2. 设 R 和 S 分别是 X 和 Y 上的等价关系. 如果 (x', y') 和 (x'', y'') 是 $\mathrm{X} \times \mathrm{Y}$ 的元素, 用 $\mathrm{T}\{(x', y'), (x'', y'')\}$ 表示关系 $\mathrm{R}\{x', x''\}$ 与 $\mathrm{R}\{y', y''\}$ 的合取. 证明 T 是 $\mathrm{X} \times \mathrm{Y}$ 上的等价关系. 用 R 和 S 的图构造 T 的图, 并定义一个从 $\mathrm{X} \times \mathrm{Y}$ 的商集到乘积 $(\mathrm{X}/\mathrm{R}) \times (\mathrm{Y}/\mathrm{S})$ 上的典范双射.

利用这些结果指出 §4 的定理 3 是定理 2 的一个特殊情形.

3. 在带有两个直角坐标轴的平面上给定坐标分别为 (x', y') 和 (x'', y'') 的两个点 P' 和 P'', 用 $\mathrm{R}\{\mathrm{P}', \mathrm{P}''\}$ 表示关系 $x'y' = x''y''$. 指出这是平面上的一个等价关系, 并且构造它的等价类.

用 $\mathrm{S}\{\mathrm{P}', \mathrm{P}''\}$ 表示关系

$$(x'y' = x''y'') \quad 和 \quad (x'x'' \geqslant 0).$$

这还是一个等价关系吗?

4. 设 A 是一个集合, 而 B 是 A 的一个子集. 用 $\mathrm{R}\{\mathrm{X}, \mathrm{Y}\}$ 表示关系 $\mathrm{X} \cap \mathrm{B} = \mathrm{Y} \cap \mathrm{B}$. 证明这是集合 $\mathscr{P}(\mathrm{A})$ 上的一个等价关系, 并且构造一个从 $\mathscr{P}(\mathrm{A})/\mathrm{R}$ 到集合 $\mathscr{P}(\mathrm{B})$ 上的双射.

5. 构造模 17 整数的加法表和乘法表.

6. 设 E 是通常空间 (看作点的集合). 选定 E 中一个点 O, 给定点 $\mathrm{P}', \mathrm{P}''$, 用 $\mathrm{R}\{\mathrm{P}', \mathrm{P}''\}$ 表示关系

$$点\ \mathrm{O}, \mathrm{P}', \mathrm{P}''\ 共线.$$

这是 E 上的一个等价关系吗? 用 E^* 表示 O 以外的点 $\mathrm{P} \in \mathrm{E}$, 即 $\mathrm{E}^* = \mathrm{E} - \{\mathrm{O}\}$. 证明 R 是 E^* 上的等价关系, 并且确定对应的等价类 (商集 E^*/R 称为**射影平面**).

7. 设 X 是所有从 $\mathbf{R}$ 到 $\mathbf{R}$ 内的映射 (对于任意 t 有定义的、实变量 t 的取实值的函数) 的集合. 给定 X 的两个元素, 用 $\mathrm{R}\{x, y\}$ 表示关系

$$存在一个数\ c > 0,\ 使得对于\ |t| < c\ 有\ x(t) = y(t).$$

证明 R 是 X 上的一个等价关系.

¶8. 设 X 是所有从 $\mathbf{R}$ 到 $\mathbf{R}$ 内的映射的集合, 选定一个整数 $n \geqslant 0$. 给定两个函数 $x, y \in \mathrm{X}$, 用 $\mathrm{R}\{x, y\}$ 表示关系

$$\lim_{t \to 0} \frac{x(t) - y(t)}{t^n} = 0$$

(通常写成形式

$$x(t) - y(t) = o(t^n),\ 当\ t \to 0).$$

证明 R 是 X 上的一个等价关系.

9. 设 X 和 Y 是两个集合, R 和 S 分别是 X 和 Y 上的等价关系, 而 f 是从 X 到 Y 内的一个映射. 考虑图表

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \downarrow p & & \downarrow q \\ X/R & & Y/S \end{array}$$

其中 p, q 分别是 X, Y 到其商集上的典范映射. 证明以下两个性质是等价的:

(i) 存在一个映射

$$\bar{f}: X/R \to Y/S,$$

使得

$$\bar{f} \circ p = q \circ f;$$

(ii) 对于任意 $x', x'' \in X$, 关系

$$x' \equiv x''(\text{mod R}) \quad 蕴含 \quad f(x') \equiv f(x'')(\text{mod S}).$$

还要证明如果条件 (ii) 满足, 则满足 (i) 的 $\bar{f}$ 是唯一的.

例子: 取 $X = Y = \mathbf{Z}$, $\mathbf{Z}$ 是有理整数集, 取 R 为关系模 r 同余, 而 S 为关系模 s 同余 (r 和 s 为给定的正整数), 最后取 f 为从 X 到 Y 内的恒等映射. 在什么情形下前面的结果可以应用?

10. 设 K 是一个单纯图 (§3, 习题 5), 假定 K 的所有一个元素的子集都是 K 的一个单形. 给定 K 的两个元素 x 和 y, 用 $R\{x, y\}$ 表示下列关系: 存在一个整数 $n \geqslant 0$ 和 K 的顶点

$$z_0 = x, z_1, \cdots, z_n = y,$$

使得对于满足条件 $0 \leqslant i < n$ 的所有 i, 集合 $\{z_i, z_{i+1}\}$ 是 K 的一个单形. 证明 R 是集合 K 上的一个等价关系. [模 R 的类称为单纯图 K 的一个**连通分支**; 称 K 是**连通的**, 如果 K 具有唯一的连通分支.]

§5　有限集和自然数

不用自然数理论, 在数学上要做任何事情显然都是不可能的, 而在集合论创立以前, 人们认为自然数理论是数学的出发点: “上帝给了我们自然数, 所有其他皆是人类的工作”, Kronnecker 如是说.

事实上, 现今借助集合论可以构造自然数, 而不必祈求任何神灵. 基本的思想乃是我们整天教育孩子们的, 自然数是用来 “数”“有限” 集合的元素的, 并且, 两个有限集有同样数目的元素, 当且仅当存在从第一个集合到第二个集合上的一个**双射** (在初等教育中人们不说 “双射”, 人们把苹果排列在梨的下方使得每个梨对应一个苹果, 并且反之亦然, 这是构造双射的一个准确、简单而直观的方法). 换句话说, 自然数概念依存于有限集概念, 并且认为在 §4, 例 3 的意义下等势的集合是相等的.

在创立集合论的初始, Cantor 立即把主要兴趣放在这个计数问题上, 不过是从更一般和更困难的角度, 即 “数” 任意有限或无限集合的元素, 这引导出超限数, 这种数可以用来区分无限集合种类的差别. 这些 “数” 在一般情形下鲜有使用, 但是在合理地陈述普通的自然数理论时不谈论这些其他的数是很困难的.

本节叙述的大部分结果没有证明, 为的是缩减篇幅 (并且因为一些定理虽然表述十分简单, 而只可惜证明是困难的, 即使在自然数理论的 “初等” 范围内). 本节仅是一个概述, 如果愿意, 可以在其中插入自然数的完备且严格的研究.

最后注意这里的自然数涉及的是在 §0 意义下的数学对象 (它们的完整定义涉及 §0, 第 9 小节的符号 τ 和 Ω), 它们用作 “具体的” 自然数的抽象模型, 不应当把二者混淆.

1. 等势集

我们回忆 (§4, 例 3) 定义: 一个集合 X 等势于集合 Y, 如果存在一个从 X 到 Y 上的双射. 关系

$$\mathrm{X} \text{ 等势于 } \mathrm{Y}$$

有时记作

$$\mathrm{Eq}(\mathrm{X}, \mathrm{Y}),$$

这是一个等价关系.

这个关系具有几个针对集合论运算的简单性质.

设 $\mathrm{X}, \mathrm{Y}, \mathrm{X}', \mathrm{Y}'$ 是四个集合, 假定 X 等势于 Y, 并且 X' 等势于 Y'. 则集合 $\mathrm{X} \times \mathrm{X}'$ 等势于集合 $\mathrm{Y} \times \mathrm{Y}'$ (如果 f 和 f' 分别是从 X 到 Y 上的和从 X' 到 Y' 上的双射, 则 $f \times f'$ 是从 $\mathrm{X} \times \mathrm{X}'$ 到 $\mathrm{Y} \times \mathrm{Y}'$ 上的一个双射). 同样从 X' 到 X 内的映射的集合 (§2 注 3)

$$\mathrm{X}^{\mathrm{X}'}$$

等势于从 Y' 到 Y 内的映射的集合 $\mathrm{Y}^{\mathrm{Y}'}$. 最后, 只要 X 和 X' 是不交的, 同时 Y 和 Y' 是不交的, 则 $\mathrm{X} \cup \mathrm{X}'$ 等势于 $\mathrm{Y} \cup \mathrm{Y}'$: 如果 f 和 f' 分别是从 X 到 Y 上的和从 X' 到 Y' 上的双射, 令

$$g(x) = \begin{cases} f(x), & \text{若 } x \in \mathrm{X}, \\ f'(x), & \text{若 } x \in \mathrm{X}', \end{cases}$$

则 g 是从 $\mathrm{X} \cup \mathrm{X}'$ 到 $\mathrm{Y} \cup \mathrm{Y}'$ 上的一个双射.

除了已经叙述的这些简单且可以直接证明的性质之外, 还有一些性质, 其叙述同样十分简单, 但是其证明困难无比. 最令人诧异的是下面一个, 其直观表述是: 给定任意两个集合 X 和 Y, 总能够 “比较” X 的元素 “个数” 和 Y 的元素 “个数”:

定理 1　给定两个集合, 下列两个断言至少有一个是真的:

X 等势于 Y 的一个子集,

Y 等势于 X 的一个子集,

并且, 如果这两个断言同时是真的, 则 X 和 Y 是等势的.

虽然 Cantor 在他最初的研究中就已经猜测到这里的结论, 但是其第二部分在 1897 年才由 Bernstein 证明, 而困难得多的第一部分到 1904 年才由 Zermelo 证明. 在本节的习题 5 中可以找到 Bernstein 的推理. 我们注意到即使在 X 和 Y 是有限集合的情形, 定理 1 也不是平凡的 (尽管直观上是显然的); 强烈建议读者深思并确认这一事实 (事先声明, 说的是证明, 而不仅仅是马马虎虎说得过去的解释).

2. 集合的基数

为了推广 "有限" 集合的元素的 "个数" 概念, Georg Cantor 引领我们给每个集合 X 指定一个新的数学对象, 记作

$$\mathrm{Card}(X),$$

称为 X 的**基数**或**势**, 其定义是下列条件满足: 两个集合 X 和 Y 是等势的, 必须且只需 $\mathrm{Card}(X) = \mathrm{Card}(Y)$.



¶¶注 1　如果存在一个集合 Ω, 其元素是所有集合, 只需取 X 对于等价关系 $\mathrm{Eq}(X, Y)$ 的类作为 $\mathrm{Card}(X)$. 但是我们知道 (§1, 注 5) 一个这样的集合 Ω 不存在, 以致这里不能使用 §4 的构造. 事实上, 一个集合 X 的基数由关系

$$\mathrm{Card}(X) = \tau_Y \mathrm{Eq}(X, Y)$$

定义, 这里利用了 §0 第 9 小节的考虑.

我们称一个数学对象 x 是一个**基数**, 如果存在一个集合 X 使得 $x = \mathrm{Card}(X)$.

对于下面出现的基数, 我们用符号 $0, 1, 2, \cdots$ 表示, 但它们自然不是 "通常的" 或 "朴素的" 数 ("通常的" 自然数是从具体经验抽象出来的概念, 而 "数学的" 自然数原则上是可以借助 §0 的过程定义的对象):

$$0 = \mathrm{Card}(\varnothing), \tag{1}$$

这是空集的基数,

$$1 = \mathrm{Card}(\{\varnothing\}), \tag{2}$$

这是仅有一个元素 $\varnothing$ 的集合的基数,

$$2 = \mathrm{Card}(\{\varnothing, \{\varnothing\}\}), \tag{3}$$

这是一个集合的基数, 其仅有的元素是空集 $\varnothing$ 和仅有一个元素 $\varnothing$ 的集合 $\{\varnothing\}$, 等等. 给定一个集合 X, 说 $\mathrm{Card}(\mathrm{X}) = 0$, 意指 X 是空集; 说 $\mathrm{Card}(\mathrm{X}) = 1$, 意指 X 是由一个元素组成的集合 (即 X 不是空集, 而且关系 $x \in \mathrm{X}$ 和 $y \in \mathrm{X}$ 蕴含 $x = y$); 说 $\mathrm{Card}(\mathrm{X}) = 2$, 意指 X 是两个元素的一个集合 (即存在 $x \in \mathrm{X}$ 和 $y \in \mathrm{X}$ 使得 $x \neq y$, 并且 $z \in \mathrm{Z}$ 蕴含 $z = x$ 或 $z = y$), 等等. 后面 (第 4 小节) 将回到这些特殊的基数.

设 x 和 y 是两个基数, 我们记

$$x \leqslant y,$$

如果存在集合 X 和 Y, 使得 $x = \mathrm{Card}(\mathrm{X}), y = \mathrm{Card}(\mathrm{Y})$, 并且 X 等势于 Y 的一个子集 —— 如果对于 X 和 Y 的一个特殊选择是这样, 那么显然对于任何其他的选择也如此. 定理 1 表明对于任何基数 x 和 y, 总有

$$x \leqslant y \quad \text{或} \quad y \leqslant x; \tag{4}$$

还有

$$x \leqslant y \text{ 和 } y \leqslant x \quad \text{蕴含} \quad x = y. \tag{5}$$

显然如果 x, y, z 是三个基数, 则

$$x \leqslant y \text{ 和 } y \leqslant z \quad \text{蕴含} \quad x \leqslant z. \tag{6}$$

这是因为, 如果存在从集合 X 到集合 Y 内的一个单射 f 和从 Y 到集合 Z 内的一个单射 g, 则存在从 X 到 Z 的一个单射, 即 $g \circ f$.

基数之间的关系 $x \leqslant y$ 的主要性质体现在下列定理中 (我们承认它而不予证明).

定理 2 设 E 是基数的一个集合, 则存在一个且仅一个基数 a 具有下列性质:

(i) 对于所有 $x \in \mathrm{E}$ 有 $x \leqslant a$ (对应的 $x \geqslant a$);

(ii) 如果一个基数 b 满足: 对于所有 $x \in \mathrm{E}$ 有 $x \leqslant b$ (对应的 $x \geqslant b$), 则有 $b \geqslant a$ (对应的 $b \leqslant a$).

这个定理表明当我们有基数的一个集合 E, 则存在大于或等于 (对应的, 小于或等于) E 的所有元素的基数, 并且在所有的大于或等于 (对应的, 小于或等于) E 的所有元素的基数当中, 存在一个小于或等于 (对应的, 大于或等于) 所有其他的基数的基数. 因而这是对于所有 $x \in \mathrm{E}$ 满足 $a \geqslant x$ (对应的, $a \leqslant x$) 的最小 (对应的, 最大) 的基数 a. 称这个基数 a 为集合 E 的**上确界** (对应的, **下确界**). 一般用记号

$$\sup(\mathrm{E}) \ (\text{对应的}, \inf(\mathrm{E}))$$

或类似记号表示它们.

注 2 在前面陈述中字母 E 表示基数的一个集合, 这是最本质的, 因为不存在含有所有基数的集合 (同样不存在含有所有集合的集合). 此外理由是, 如果



存在一个这样的集合, 定理 2 应用到这个集合就得到存在一个基数 a, 大于或等于所有基数 —— 而这是不可能的, 因为可以证明对于所有基数 a, 我们有严格不等式

$$a < 2^a.$$

这里我们还发现, 如果赋予 "集合" 一词以直观意义, 就会显示出逻辑矛盾.

还要注意, 如果 E 是基数的一个集合, 基数 sup(E) 和 inf(E) 不必属于 E. 举例来说, 取 E 是所有有限基数的集合 **N** (参见后面), sup(E) 就是可数势 (第 5 小节), 于是不在 E 内.

不过当 E 是有限基数即自然数 (这个概念将在第 4 小节定义) 的一个集合时, 总有 $\inf(\mathrm{E}) \in \mathrm{E}$. 事实上, 由于 inf(E) 小于或等于所有 $x \in \mathrm{E}$, 显然 inf(E) 是有限的; 如果 inf(E) 不属于 E, 将有 $\inf(\mathrm{E}) < x$, 于是对于所有 $x \in \mathrm{E}$ 将有

$$\inf(\mathrm{E}) + 1 \leqslant x.$$

再根据定理 2 的性质 (ii) 将有

$$\inf(\mathrm{E}) + 1 \leqslant \inf(\mathrm{E}),$$

这是不可能的, 因为 inf(E) 是有限的.

换句话说, *如果 E 是自然数的一个集合, 则存在 E 的一个元素, 它小于或等于所有其他的元素*, 这是在实际中经常使用的一个结果.

3. 基数的运算

设 x 和 y 是两个基数, 并令 $x = \mathrm{Card}(\mathrm{X})$, $y = \mathrm{Card}(\mathrm{Y})$, 则称基数

$$xy = \mathrm{Card}(\mathrm{X} \times \mathrm{Y}) \tag{7}$$

为 x 乘以 y 的**乘积**. 显然把 X (对应的, Y) 代以等势于 X (对应的, Y) 的集合, 乘积不改变.

这个运算满足以下等式:

$$xy = yx; \quad x(yz) = (xy)z, \quad 0x = 0; \quad 1x = x. \tag{8}$$

比如为了证明第二个等式, 只需注意到, 如果 X, Y 和 Z 是三个集合, 则有 $\mathrm{X} \times (\mathrm{Y} \times \mathrm{Z})$ 等势于 $(\mathrm{X} \times \mathrm{Y}) \times \mathrm{Z}$, 这是显然的, 如果令后者的每个元素 $((a, b), c)$ 对应前者的元素 $(a, (b, c))$.



¶注 3 虽然直观看似乎成立, 而实际上

$$xz = yz \quad \text{蕴含} \quad x = y$$

是错误的, 即使 $z \neq 0$. 为了得到一个反例, 承认集合 $\mathbf{N}$ 的存在性 (见后面), 其元素是自然数 $0, 1, 2, \cdots$, 取 $z = \mathrm{Card}(\mathbf{N})$, $x = 1, y = 2$, x 是只有一个元素 a 的集合 X 的基数, y 是有两个元素 b, c 的集合 Y 的基数, 所有的事情归结为构造从 $\mathrm{X} \times \mathbf{N}$ 到 $\mathrm{Y} \times \mathbf{N}$ 上的一个双射 f,

$$f(a, n) = \begin{cases} (b, p), & \text{若 } n = 2p, \\ (c, p), & \text{若 } n = 2p + 1. \end{cases}$$

显然如果 $z \neq 0$, 而 x, y, z 是自然数 (见后面) 那么

$$\text{关系 } xz = yz \text{ 蕴含 } x = y$$

是成立的.

设 x 和 y 是两个基数, 选择两个不交的集合 X 和 Y, 使得 $x = \mathrm{Card(X)}$, $y = \mathrm{Card(Y)}$, 则称基数

$$x + y = \mathrm{Card(X \cup Y)} \ (\text{对于 } \mathrm{X \cap Y} = \varnothing) \tag{9}$$

为 x 与 y 的**和**. 立刻就能够验证它不依赖 X 和 Y 的选取. 下列等式成立:

$$x + y = y + x, \quad x + (y + z) = (x + y) + z, \quad 0 + x = x. \tag{10}$$

比如最后一个等式表示对于所有的集合 X, X 等势于 $\mathrm{X} \cup \varnothing$, 这不足为奇, 其实甚至有

$$\mathrm{X} = \mathrm{X} \cup \varnothing.$$

注 4 上面所给出的 $x + y$ 的定义假定了总可以找到两个不交的集合 X 和 Y 使得 $x = \mathrm{Card(X)}$, $y = \mathrm{Card(Y)}$. 为此令 $x = \mathrm{Card(X')}$, $y = \mathrm{Card(Y')}$, 取

$$\mathrm{X} = \mathrm{X'} \times \{a\}, \quad \mathrm{Y} = \mathrm{Y'} \times \{b\},$$

其中 $a \neq b$, 那么 X 和 Y 分别等势于 X′ 和 Y′, 并且是不交的, 因为关系 $(x, a) = (y, b)$ 必然导致 $a = b$.

¶注 5 这里



$$\text{关系 } x + z = y + z \text{ 蕴含 } x = y$$

也是错误的. 例如可以同时成立

$$x + x = x \quad \text{和} \quad x \neq 0.$$

为此取 $x = \mathrm{Card}(\mathbf{N})$, 其中 $\mathbf{N}$ 是自然数 $0, 1, 2, \cdots$ 的集合 (见后面), 那么有 $x + x = \mathrm{Card(Y)}$, Y 是序偶 (n, u) 的集合, $n \in \mathbf{N}$, 而 $u = 0$ 或 $u = 1$; 令

$$f(n, u) = \begin{cases} 2n, & \text{如果 } u = 0, \\ 2n + 1, & \text{如果 } u = 1, \end{cases}$$

我们就得到一个从 Y 到 **N** 上的双射 f, 这表明了在这种情形有 $x+x=x$.

在加法和乘法之间还有 “分配” 关系

$$x(y+z)=xy+xz, \tag{11}$$

它表示给定了三个集合 X, Y, Z, 那么 $\mathrm{X}\times(\mathrm{Y}\cup\mathrm{Z})$ 等势于 $(\mathrm{X}\times\mathrm{Y})\cup(\mathrm{X}\times\mathrm{Z})$, 而这在几何上是显然的.

最后考虑两个基数 x 和 y, 并且令 $x=\mathrm{Card}(\mathrm{X}), y=\mathrm{Card}(\mathrm{Y})$. 定义从 Y 到 X 内的所有映射的集合的基数

$$x^y=\mathrm{Card}(\mathrm{X}^{\mathrm{Y}}). \tag{12}$$

这个运算称为基数的**取幂**, 它满足以下等式:

$$x^{y+z}=x^y\cdot x^z;\quad (xy)^z=x^zy^z;\quad (x^y)^z=x^{yz};\quad x^0=1;\quad x^1=x. \tag{13}$$

例如, 为了证明第一个等式, 令 $x=\mathrm{Card}(\mathrm{X}), y=\mathrm{Card}(\mathrm{Y}), z=\mathrm{Card}(\mathrm{Z})$, 并且假定 Y 和 Z 是不交的, 由此得 $y+z=\mathrm{Card}(\mathrm{Y}\cup\mathrm{Z})$. 那么现在需要证明集合

$$\mathrm{X}^{\mathrm{Y}\cup\mathrm{Z}}\quad 和 \quad \mathrm{X}^{\mathrm{Y}}\times\mathrm{X}^{\mathrm{Z}}$$

是等势的, 即构造从第一个集合到第二个集合上的一个双射 f. 为此, 设 u 是第一个集合的一个元素, 即从 $\mathrm{Y}\cup\mathrm{Z}$ 到 X 上的一个映射, 设 u_{Y} 和 u_{Z} 是 u 分别在 Y 和 Z 上的限制, 令 $f(u)=(u_{\mathrm{Y}},u_{\mathrm{Z}})$, 读者 (利用 Y 和 Z 是不交的假设) 容易验证 f 是双射.

注 6 考虑两个元素的一个集合, 比如 $\mathrm{A}=\{0,1\}$, 设 f 是从集合 X 到 A 内的一个映射, 显然 f 由使得 $f(x)=0$ 的 x 集合

$$f^{-1}(0)\subset\mathrm{X}$$

唯一决定. 立即就可验证如此定义的从 A^{X} 到 $\mathscr{P}(\mathrm{X})$ 的映射 $f\to f^{-1}(0)$ 是双射. 由此推知

$$\mathrm{Card}(\mathscr{P}(\mathrm{X}))=2^{\mathrm{Card}(\mathrm{X})}.$$



注 7 可以证明对于所有基数 x 有



$$x<2^x$$

(即 $x\leqslant 2^x$, 而且 $x\neq 2^x$). 特别的, 如果 X 是一个集合, 那么 $\mathscr{P}(\mathrm{X})$ 不等势于 X. 见习题 1.

¶**注 8** 不仅可以定义两个基数的和与乘积, 还可以更一般地定义基数的一个任意 (甚至无穷的) 族的和与乘积. 定义如下进行.

首先设 $(\mathrm{X}_i)_{i\in\mathrm{I}}$ 是一个集族, 一个集合, 其元素是所有的族 $(x_i)_{i\in\mathrm{I}}$, 对于所有 $i \in \mathrm{I}$ 有 $x_i \in \mathrm{X}_i$, 称为 X_i 的笛卡儿乘积, 记作

$$\prod_{i\in\mathrm{I}} \mathrm{X}_i.$$

这个概念显然推广了 §2 第 2 小节里的概念. 因为如果 $\mathrm{I} = \{1,2\}$, 并且把集族 $(x_i)_{i\in\mathrm{I}}$ 和序偶 (x_1, x_2) 等同, 就可以把 $\mathrm{X}_1, \mathrm{X}_2$ 的乘积和集合 $\mathrm{X}_1 \times \mathrm{X}_2$ 等同.

交代好了这些, 设 $(x_i)_{i\in\mathrm{I}}$ 是基数的任意一个族, 对于每个 $i \in \mathrm{I}$ 选择一个集合 X_i 使得 $x_i = \mathrm{Card}(\mathrm{X}_i)$, 定义

$$\prod_{i\in\mathrm{I}} x_i = \mathrm{Card}\left(\prod_{i\in\mathrm{I}} \mathrm{X}_i\right).$$

这样定义的基数称为**基数族** $(x_i)_{i\in\mathrm{I}}$ **的乘积**.

同样由关系

$$\sum_{i\in\mathrm{I}} x_i = \mathrm{Card}\left(\bigcup_{i\in\mathrm{I}} \mathrm{X}_i\right)$$

定义**基数族** $(x_i)_{i\in\mathrm{I}}$ **的和**, 显然要选择 X_i 使得它们两两是不交的. 作为例子, 设对于每个 $i \in \mathrm{I}$ 有 $x_i = 1$, 我们可以对于每个 $i \in \mathrm{I}$ 取 $\mathrm{X}_i = \{i\}$, 这些集合的并正好是 I, 则有

$$\mathrm{Card}(\mathrm{I}) = \sum_{i\in\mathrm{I}} x_i, \text{ 其中对于每个 } i \in \mathrm{I} \text{ 有 } x_i = 1. \tag{14}$$

直观地说, 这表明所有的基数是其每个项 (一般有无限多个) 都等于 1 的一个和, 对于通常的自然数, 这个性质是人所共知的.

4. 有限集和自然数

以下结果是基本的:

定理 3 *设 X 是一个集合, 以下性质是等价的:*

a) 包含于 X 且与 X 等势的集合只有 X 本身;

b) $\mathrm{Card}(\mathrm{X}) \neq \mathrm{Card}(\mathrm{X}) + 1$.

假定 $\mathrm{Card}(\mathrm{X}) = \mathrm{Card}(\mathrm{X}) + 1$. 用 a 表示一个不属于 X 的元素, 于是存在从 $\mathrm{X} \cup \{a\}$ 到 X 上的一个双射 f. X 在 f 下的像显然等势于 X 并且严格包含于 X.

反之假定 X 等势于严格包含于 X 内的一个集合 X′. 于是由于 $\mathrm{X} = \mathrm{X}' \cup (\mathrm{X} - \mathrm{X}')$, 我们有

$$\mathrm{Card}(\mathrm{X}) = \mathrm{Card}(\mathrm{X}') + \mathrm{Card}(\mathrm{X} - \mathrm{X}').$$

但是由于 $X - X'$ 是非空的, $\mathrm{Card}(X - X') \geqslant 1$, 故

$$\mathrm{Card}(X) \geqslant \mathrm{Card}(X) + 1 \geqslant \mathrm{Card}(X),$$

根据定理 1 得 $\mathrm{Card}(X) = \mathrm{Card}(X) + 1$, 这就完成了定理 3 的证明.

我们说一个集合 X 是**有限的**, 如果它具有上述定理中的性质 a) 和 b), 在相反的情形, 则说它是**无限的**. 同样, 如果 $x \neq x+1$, 则说基数 x 是**有限的**, 如果 $x = x+1$, 则说它是**无限的**. 一个有限的基数称为**自然数**, 一个无限的基数称为**超限数**.

自然数具有非常简单的性质 (所有的人都知道), 首先, 如果 x 和 y 是自然数, 则 xy 和 x^y 也是自然数. 更一般的, 如果 $(x_i)_{i\in I}$ 是自然数的有限族 (我们说一个族 $(x_i)_{i\in I}$ 是**有限的**, 如果其指标集 I 是有限的), 那么基数 (注 8)

$$\prod_{i\in I} x_i \quad 和 \quad \sum_{i\in I} x_i$$

也是有限的.

如果 x 是一个自然数, 则所有满足 $y \leqslant x$ 的基数也是有限的 (换句话说, 有限集的所有子集都是有限集), 并且存在一个且仅一个基数 z 使得

$$x = y + z.$$

z 是有限的, 称为 x 和 y 之间的差, 并且记为

$$z = x - y.$$

如果 $x = \mathrm{Card}(X)$, $y = \mathrm{Card}(Y)$, 并且 $Y \subset X$, 则显然有 $z = \mathrm{Card}(X - Y)$.

最后当谈论自然数时, 在注 3 和注 5 中所考察的病态情形不再出现. 精确说来有

如果 z 是有限的, 则关系 $x + z = y + z$ 蕴含 $x = y$,

同样有

如果 z 是有限的并且不是零, 则关系 $xz = yz$ 蕴含 $x = y$.

不言而喻, 上面定义的基数 $0, 1, 2, \cdots$ 是有限的.

经常需要下列性质:

定理 4 设 X 是一个有限集合, f 是从 X 到 X 内的一个映射, 则以下性质是等价的:

a) f 是单射的;

b) f 是满射的;

c) f 是双射的.

显然只需证明 a) 和 b) 的等价性. 设 f 是单射的, f 是从 X 到 X 的一个子集 $f(\mathrm{X})$ 上的双射. 因此 $f(\mathrm{X})$ 等势于 X, 因为 X 是有限的, 故 $f(\mathrm{X}) = \mathrm{X}$, 随之 a) 蕴含 b).

假定 f 是满射的, 那么存在 (§2, 定理 4) 从 X 到 X 内的映射 h 使得 $f \circ h = j_{\mathrm{X}}$, 后者是恒等映射; h 显然是单射, 那么根据刚证明的 a) 蕴含 b), h 是满射的, 随之是双射的, 从而 f 是 h 的反映射, 因此是单射的, 定理得证.

5. 自然数集合 N

前面的考虑使得有限集的存在性变得显然, 因为

$$\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}, \cdots$$

都是有限集 —— 正是在这里, 数学家获得了某种实实在在的结果, 尽管一切都建立在空集概念的基础之上.

反之, *无限集的存在性*不是明显的; 事实上尽管如此, 这是一个数学公理. 那些认为这个观点令人吃惊和违反直观的读者, 应当好好回忆一下, 在数学里, 人们坚持逻辑地证明断言, 特别强调, "存在" 这个词跟物理学和神学里的意义根本不一样.



定理 5 *设 X 是一个无限集, 则所有有限集等势于 X 的一个子集.*

所有的事情都归结为证明: 对于所有的有限基数 y 和所有的无限基数 x 都有 $y \leqslant x$. 而如果这不成立, 将有 $y \geqslant x$, 那么像从前一小节所看到的将有 x 是有限的.

定理 6 *存在一个且仅一个集合 $\mathbf{N}$, 使得关系*

$$x \in \mathbf{N}$$

等价于关系

$$x \text{ 是一个自然数.}$$

集合 $\mathbf{N}$ 是无限的.

根据 §1 的定理 2, $\mathbf{N}$ 的唯一性是显然的. $\mathbf{N}$ 的存在性来自以下事实: 如果选择一个无限基数 a (根据无限集的存在性这是可能的), 那么正如刚刚看到的, 所有的自然数 x 满足关系 $x < a$. 于是只需证明, 对于所有基数 a, 满足 $x < a$ 的基数是一个集合的元素, 我们承认这个断言(*).

最后设 $\mathbf{N}$ 是有限的, 对于每个 $n \in \mathbf{N}$, 选择一个集合 X_n 使得 $\mathrm{Card}(\mathrm{X}_n) = n$, 因为 $\mathbf{N}$ 和每个 X_n 是有限的, 集合

$$\mathrm{X} = \bigcup_{n \in \mathbf{N}} \mathrm{X}_n$$

(*) 令 $a = \mathrm{Card}(\mathrm{A})$, $x \leqslant a$ "双射对应" 于第 1 小节中的关系 $\mathrm{Eq}(\mathrm{X}, \mathrm{Y})$ 在 $\mathscr{P}(\mathrm{A})$ 内的等价类; 这就 (似乎) 说明为什么这些 x 在一个集合里.

也是有限的, 而由于 X_n 包含于 X 内, 我们看到, 如果 **N** 是有限的, 就存在一个自然数 $x = \mathrm{Card}(X)$ 使得对于所有有限的 n 有 $n \leqslant x$; 此时由于 x 是有限的, $x+1$ 也是有限的, 故有 $x+1 \leqslant x$, 但还有 $x \leqslant x+1$, 故有 $x = x+1$, 这与 x 是有限的相矛盾, 这就完成了证明 (并不严格).

概言之, 我们看到以下两个断言:

存在一个无限集;

存在一个集合, 其元素是所有自然数,

是等价的.

基数

$$\mathrm{Card}(\mathbf{N})$$

称为可数势, 如果一个集合等势于 **N**, 换句话说, 存在一个从自然数集合到 X 的元素的集合上的双射

$$n \to x_n,$$

则说它是可数的,

例 1 非负分数 (假定这个概念已经定义) 的集合是可数的. 为了确信这一事实, 只需指出可以书写非负分数 (它本来依赖两个自然数) 成一个无穷序列, 含有这些数一次且仅一次. 如下进行:

$$\frac{0}{1};\quad \frac{1}{1};\quad \frac{1}{2},\frac{2}{1};\quad \frac{1}{3},\frac{3}{1};\quad \frac{1}{4},\frac{2}{3},\frac{3}{2},\frac{4}{1};\quad \frac{1}{5},\frac{5}{1};\quad \cdots.$$

首先写满足 $p+q=1$ 的数 p/q, 然后写使得 $p+q=2$ 并且没有写过的数, 然后写使得 $p+q=3$ 并且没有写过的数, 如此下去.

不存在比自然数更多的非负分数这一点初看违背常识, 但是在数学中让 "常识" 失去信誉正是 Cantor 最重要的成功之一.



注 9 根据注 7, $\mathscr{P}(\mathbf{N})$ 的基数

$$2^{\mathrm{Card}(\mathbf{N})}$$

严格大于 $\mathrm{Card}(\mathbf{N})$, 称为**连续统势**. 可以证明

$$2^{\mathrm{Card}(\mathbf{N})} = \mathrm{Card}(\mathbf{R}),$$

其中 **R** 是实数集 (或直线的点的集).

自从 20 世纪初, 人们徒劳无益地努力证明连续统假设, 即 **R** 的所有无限子集等势于 **N** 或 **R**. 在 1939 年人们 (K. Gödel) 可以证明, 像选择公理一样 (见 §2, 第 8 小节, 注 8), 连续统假设跟集合论的其他公理是可相容的. 更近些, 人

们 (P. Cohen, 1963) 还证明了连续统假设和和选择公理是相互独立的断言 (换句话说, 集合论的公理, 含有选择公理, 不蕴含连续统假设, 反之, 连续统假设也不蕴含选择公理). 关于这个问题, 在 P. Cohen 的著作《集合论和连续统假设》(Benjamin, New York, 1966) 里有漂亮的陈述. 请注意, 尽管有了这些令人印象深刻的结果, 集合论的无矛盾问题仍然是没有解决的.

6. 数学归纳法推理

数学归纳法推理建立在下列断言的基础之上:

定理 7 设 $\mathrm{R}\{n\}$ 是含有一个变量 $n \in \mathbf{N}$ 的关系. 假定关系 $\mathrm{R}\{0\}$ 成立, 并且对于所有 $n \in \mathbf{N}$

$$\text{关系 } \mathrm{R}\{n\} \text{ 蕴含 } \mathrm{R}\{n+1\}$$

成立, 则关系 $\mathrm{R}\{n\}$ 对于所有 $n \in \mathbf{N}$ 成立.

事实上, 设 E 是使得 $\mathrm{R}\{n\}$ 成立的 $n \in \mathbf{N}$ 的集合, 应当证明 $\mathrm{E} = \mathbf{N}$; 或同样的, 证明 $\mathrm{F} = \mathbf{N} - \mathrm{E}$ 是空集. 假定 F 不是空集, 那么 (注 2) 存在一个 $a \in \mathrm{F}$ 使得对于所有 $n \in \mathrm{F}$ 有 $a \leqslant n$; 而由假设 $0 \in \mathrm{E}$, 因此 $a \geqslant 1$; 于是对于一个 $n \in \mathbf{N}$ 有 $a = n+1$, 而由于 $n < a$, 不可能有 $n \in \mathrm{F}$, 从而 $\mathrm{R}\{n\}$ 成立; 但是根据假设 $\mathrm{R}\{n\}$ 蕴含 $\mathrm{R}\{n+1\}$, 故 $\mathrm{R}\{n+1\}$ 即 $\mathrm{R}\{a\}$ 成立, 这与 $a \in \mathrm{F}$ 的事实相矛盾. 故得定理.

在实际中人们经常利用定理 6 的变种, 特别的一个变种是把 $\mathrm{R}\{n\}$ 蕴含 $\mathrm{R}\{n+1\}$ 替换为下列形式:

$$\text{关系 } \mathrm{R}\{1\}, \cdots, \mathrm{R}\{n\} \text{ 的合取蕴含 } \mathrm{R}\{n+1\}$$

(常常为了证明 $\mathrm{R}\{n+1\}$, 不仅要使用 $\mathrm{R}\{n\}$, 而且还要使用 $\mathrm{R}\{n+1\}$ 前面的所有断言).

例 2 对于每个自然数 n, 用 E_n 表示使得 $x \leqslant n$ 的自然数 x 的集合; 我们要对 n 用数学归纳法证明对于所有的 n

$$\mathrm{Card}(\mathrm{E}_n) = n + 1.$$

事实上, 对于 $n = 0$ 我们有 $\mathrm{E}_n = \{0\}$, 因此 $\mathrm{Card}(\mathrm{E}_0) = 1$. 余下的是证明对于自然数 n 的断言蕴含对于自然数 $n+1$ 的断言. 设 $x \leqslant n+1$, 那么或者 $x \leqslant n$, 从而 $x \in \mathrm{E}_n$, 或者

$$n < x \leqslant n+1,$$

从而 $x = n+1$ (见下面的说明), 于是 $\mathrm{E}_{n+1} = \mathrm{E}_n \cup \{n+1\}$, 随即得到

$$\mathrm{Card}(\mathrm{E}_{n+1}) = \mathrm{Card}(\mathrm{E}_n) + 1,$$

这显然说明关系 $\mathrm{Card}(\mathrm{E}_n) = n+1$ 蕴含 $\mathrm{Card}(\mathrm{E}_{n+1}) = n+2$. 这就证明了对于所有的 n 有 $\mathrm{Card}(\mathrm{E}_n) = n+1$.

上面使用了下列事实:

$$\text{关系 } n < x \leqslant n+1 \text{ 蕴含 } x = n+1.$$

可以如下证明这一结论. 设 A 是一个集合满足

$$\mathrm{Card}(\mathrm{A}) = n+1,$$

由于 $x \leqslant n+1$, 存在一个集合 X 使得 $x = \mathrm{Card}(\mathrm{X})$, 并且 $\mathrm{X} \subset \mathrm{A}$. 由于 $n < x$, 存在一个集合 B 使得

$$n = \mathrm{Card}(\mathrm{B}), \quad \mathrm{B} \subset \mathrm{X}, \quad \mathrm{B} \neq \mathrm{X}.$$

我们有

$$n+1 = \mathrm{Card}(\mathrm{A}) = \mathrm{Card}(\mathrm{B}) + \mathrm{Card}(\mathrm{A}-\mathrm{B}) = n + \mathrm{Card}(\mathrm{A}-\mathrm{B}),$$

于是 $\mathrm{Card}(\mathrm{A}-\mathrm{B}) = 1$, 这说明 B 在 A 内的补集含有一个元素. 由于 X 包含 B 并且不同于 B, 故 $\mathrm{X} = \mathrm{A}$, 随之得到所要求的 $x = n+1$.

7. 组合分析

下面, 给定一个自然数 n, 称所有使得 $\mathrm{Card}(\mathrm{X}) = n$ 的集合 X 为 n 元集.

定理 8 (牧羊人原理) 设 f 是从集合 X 到一个集合 Y 上的映射. 假定 Y 是一个 q 元集, 并且对于所有 $y \in \mathrm{Y}$, 集合 $f^{-1}(y) \subset \mathrm{X}$ 是 p 元集, 则 X 是一个 pq 元集.

取定一个 p 元集 F, 对于每个 $y \in \mathrm{Y}$, 取定一个从 F 到 $f^{-1}(y)$ 上的一个映射 u_y. 令

$$u(y,z) = u_y(z) \quad \text{对于 } y \in \mathrm{Y} \text{ 和 } z \in \mathrm{F},$$

这样就定义了映射

$$u : \mathrm{Y} \times \mathrm{F} \to \mathrm{X}.$$

考虑到 f 是满射, 立即就可以验证 u 是一个双射. 于是

$$\mathrm{Card}(\mathrm{X}) = \mathrm{Card}(\mathrm{Y} \times \mathrm{F}) = \mathrm{Card}(\mathrm{Y}) \times \mathrm{Card}(\mathrm{F}) = qp,$$

这就完成了证明.

定理 9 设 X 是一个 p 元集, Y 是一个 q 元集, 则从 Y 到 X 内的映射的集合是 p^q 元集.

这个定理其实就是 p^q 的定义 (第 3 小节). 当然, 为了使它可用, 必需验证当涉及的是自然数时, 基数的乘幂归结为所有人所熟知的运算. 第 3 小节的公式 (13) 指出

$$n^1 = n, \quad n^0 = 1, \quad n^{p+q} = n^p n^q,$$

于是有

$$n^2 = n^{1+1} = n^1 n^1 = n \cdot n; \quad n^3 = n^{2+1} = n^2 n^1 = n \cdot n \cdot n, \ \cdots.$$

注 10 取 Y 是满足 $1 \leqslant i \leqslant q$ 的自然数 i 的集合, 从 Y 到 X 内的一个映射是 X 的元素的一个族 $(x_i)_{i \in \mathrm{Y}}$, 这样的族经常写作



$$(x_i)_{1 \leqslant i \leqslant q},$$

并且称为 **X 的 q 个元素的一个排列**(*). 定理 8 断言如果 X 是 p 元集, 则这些排列的个数是 p^q.

定理 10 设 X 是一个 p 元集, Y 是一个 q 元集, 假定 $p \leqslant q$. 则从 X 到 Y 内的单射的个数是

$$\frac{q!}{(q-p)!},$$

其中对于自然数 n,

$$\text{如果 } n \neq 0, \text{ 令 } n! = 1 \cdot 2 \cdot 3 \cdots n; \quad \text{如果 } n = 0, \text{ 令 } n! = 1.$$

如果 $p = 0$, 集合 X 是空集, 仅有一个从 X 到 Y 内的单射, 因而在这种情形定理是成立的. 剩下的 (由定理 6) 是要证明如果它对于一个自然数 p 是成立的, 则对于 $p+1$ 也是成立的.

假设 $\mathrm{Card}(\mathrm{X}) = p+1$ 和 $\mathrm{Card}(\mathrm{Y}) = q \geqslant p+1$. 选定一个 $a \in \mathrm{X}$, 并令 $\mathrm{X}' = \mathrm{X} - \{a\}$, 从而 X′ 具有 p 个元素. 把从 X 到 Y 内的单射的集合记作 I. 令

$$u(f) = f(a), \text{ 对于所有 } f \in \mathrm{I},$$

这样就定义了一个映射

$$u : \mathrm{I} \to \mathrm{Y},$$

这个从 I 到 Y 内的映射显然是满射的 (换句话说, 对于所有 $b \in \mathrm{Y}$, 存在一个单射使得 $f(a) = b$). 对于一个给定的 $b \in \mathrm{Y}$, 考虑使得 $f(a) = b$ 的 $f \in \mathrm{I}$; 令 $\mathrm{Y}' = \mathrm{Y} - \{b\}$, 一个这样的 f 显然诱导出一个从 X′ 到 Y′ 内的单射 f', 反之, 所有从 X′ 到 Y′ 内的

(*) 这里的排列中的对象允许重复. —— 译者注

单射 f' 可以补充成为一个从 X 到 Y 内的单射 f, 使得 $f(a) = b$. 我们看到, 利用归纳假设, 使得 $u(f)$ 给定的 $f \in \mathrm{I}$ 的个数是

$$\frac{(q-1)!}{[(q-1)-(p-1)]!} = \frac{(q-1)!}{(q-p)!}.$$

由于 u 映射 I 到 q 元集 Y 上, 牧羊人原理指出

$$\mathrm{Card}(I) = q \cdot \frac{(q-1)!}{(q-p)!},$$

由于

$$q! = q \cdot (q-1)!$$

证明完成.

推论 *n 元集 X 的置换的个数是 $n!$.*

一个置换是从 X 到 X 内的一个双射 (§2, 第 8 小节), 但由于 X 是有限集, X 的置换也是从 X 到 X 内的一个单射 (定理 4). 余下的则是应用定理 9, 其中取 $\mathrm{Y} = \mathrm{X}$ 和 $p = q = n$, 并且注意到 $(n-n)! = 0! = 1$.

数 $n!$ 读作 n **的阶乘**. 我们有关系

$$0! = 1, \quad 1! = 1, \quad 2! = 2, \quad 3! = 6, \quad 4! = 24, \quad 5! = 120, \quad \cdots.$$

定理 11 *设 X 是有 n 个元素的集合, 而 p 是小于或等于 n 的一个自然数. 则包含在 X 内的 p 个元素的集合的个数是*

$$\frac{n(n-1)\cdots(n-p+1)}{p!} = \frac{n!}{p!(n-p)!}.$$

选定一个有 p 个元素的集合 E. 用 I 表示从 E 到 X 内的单射的集合, 用 P 表示 X 的 p 元子集的集合.

对于所有 $f \in \mathrm{I}$, 显然 $f(\mathrm{E})$ 是 X 的一个 p 元子集. 于是令

$$u(f) = f(\mathrm{E}).$$

这就定义了一个映射

$$u : \mathrm{I} \to \mathrm{P}.$$

这个映射是满射的, 这是因为 X 的一个 p 元子集 Y 等势于 E, 故是从 E 到 X 内的一个单射 E 的像.

我们来计算使得 $f(\mathrm{E}) = \mathrm{Y}$ 的 $f \in \mathrm{I}$ 的个数, 这里 Y 是 X 的一个给定的子集. 设 f_0 是一个这样的映射, 人们把 f_0 与集合 E 的一个任意的置换复合就可以得到其他的映射: 如果 $f(\mathrm{E}) = f_0(\mathrm{E})$, 那么对于所有 $x \in \mathrm{E}$, 存在唯一的一个 $s(x) \in \mathrm{E}$ 使得

$f(x) = f_0(s(x))$, s 显然是 E 的一个置换. 于是令 E 的每个置换对应从 E 到 X 内的映射 $f = f_0 \circ s$, 我们就得到了从 E 的置换的集合到使得 $f(\mathrm{E}) = \mathrm{Y}$ 的 $f \in \mathrm{I}$ 的集合上的一个双射.

对于给定的 Y, 使得 $f(\mathrm{E}) = \mathrm{Y}$ 的 $f \in \mathrm{I}$ 的个数是 $p!$ (定理 9 的推论). 根据牧羊人原理, 我们得到

$$\mathrm{Card}(\mathrm{I}) = p!\mathrm{Card}(\mathrm{P}).$$

于是根据定理 9 有

$$\mathrm{Card}(\mathrm{P}) = \frac{\mathrm{Card}(\mathrm{I})}{p!} = \frac{n!}{p!(n-p)!},$$

这就完成了证明.

习惯上记

$$\frac{n!}{p!(n-p)!} = \binom{n}{p},$$

称这些数为**二项式系数**, 后面 (§8, 第 4 小节) 就会明白这样命名的缘由.

注 11 一个集合 X 的 p 个元素的一个子集, 曾经称为 **X 的 p 个元素的一个组合**. 传统的著作定义组合是 X 的两两不同的元素的一个序列

$$x_1, \cdots, x_p$$

并且把仅仅元素排列次序不同的两个序列视为等同. 显然使用集合论的语言更加自然.

8. 有理整数(*)

除了自然数, 数学里还需要 "任意符号" 的整数, 或**有理整数**. 我们概要地指出如何定义它们.

基本的思路是, 如果 x 和 y 是两个自然数, 存在一个有理整数 z 使得

$$x + z = y,$$

这说明正是负数的创立, 使得减法在所有情形下是可能的. 如果我们假定已经构造了有理整数集 $\mathbf{Z}$, 那么就可以用

$$\mathrm{D}(x, y) = x - y$$

(*) 这一小节的目的是验证读者业已熟悉的结果, 所以初读可以忽略.

我们指出 "任意符号" 的整数 (我们像所有数学家那样, 称为有理整数) 在过去的法国中学教学里, 称为 "代数整数" (同样称整数的或非整数的任意符号的数为 "代数数", 这种数也是像所有数学家那样, 我们称之为实数). 术语的这种分歧不会存在, 如果数学家已经在完全明确的意义下使用词组 "代数整数" 和 "代数数", 后面将会定义这些词组 (§11, 例 11, 以及 §26 的习题).

定义一个满射

$$\mathrm{D}: \mathbf{N} \times \mathbf{N} \to \mathbf{Z}.$$

显然有

$$\mathrm{D}(x, y) = \mathrm{D}(x', y') \quad \text{等价于} \quad x + y' = x' + y.$$

这些说明可以提示下面的构造 (显然在后面不再假定问题已经解决).

为了构造有理整数集 $\mathbf{Z}$, 我们以自然数的序偶的集合 $\mathbf{N} \times \mathbf{N}$ 做出发点, 在这个集合上定义一个等价关系 R, 我们说自然数的两个序偶 (x, y) 和 (x', y') 是模 R 等价的, 当且仅当

$$x + y' = x' + y.$$

R 是一个等价关系的验证是简单的, 留给读者去做. 做了验证之后, 定义

$$\mathbf{Z} = (\mathbf{N} \times \mathbf{N})/\mathrm{R},$$

称 $\mathbf{Z}$ 的所有元素为**有理整数**(*).

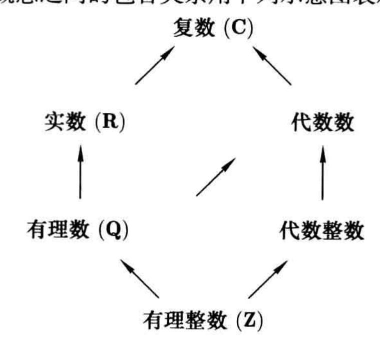
显然必须在有理整数上定义代数运算, 并且要指出如何把自然数看作特殊的有理整数. 为此记从 $\mathbf{N} \times \mathbf{N}$ 到 $\mathbf{Z}$ 上的典范映射 (§4, 第 2 小节) 为 D, 为了定义两个有理整数 z 和 z' 的**和**与**乘积**, 取 $(x, y), (x', y') \in \mathbf{N} \times \mathbf{N}$, 使得

$$z = \mathrm{D}(x, y), \quad z' = \mathrm{D}(x', y'),$$

令(**)

$$z + z' = \mathrm{D}(x + x', y + y'),$$
$$zz' = \mathrm{D}(xx' + yy', xy' + x'y);$$

(*) 最常遇到的数的各种概念之间的包含关系用下列示意图表示



其中每个箭头表示出发集包含在到达集内.

(**) 这个构造由人们希望最终实现公式

$$(x - y) + (x' - y') = (x + x') - (y + y'),$$
$$(x - y) \cdot (x' - y') = (xx' + yy') - (xy' + x'y)$$

来解释.

如果愿意, 还可以应用 §4 的定理 3: 采用定理 3 的记号, 取 $\mathrm{X} = \mathrm{Y} = \mathrm{Z} = \mathbf{N} \times \mathbf{N}$, $\mathrm{R} = \mathrm{S} = \mathrm{T}$, 取映射 $f : \mathrm{X} \times \mathrm{Y} \to \mathrm{Z}$ 或者为

$$f((x,y),(x',y')) = (x+x', y+y'),$$

或者为

$$f((x,y),(x',y')) = (xx'+yy', xy'+x'y).$$

必须验证定理 3 的条件 a), 这就是说, 应当证明关系

$$(x,y) \equiv (u,v) \pmod{\mathrm{R}} \quad \text{和} \quad (x',y') \equiv (u',v') \pmod{\mathrm{R}}$$

蕴含

$$(x+x', y+y') \equiv (u+u', v+v') \pmod{\mathrm{R}}$$

和

$$(xx'+yy', xy'+x'y) \equiv (uu'+vv', uv'+u'v) \pmod{\mathrm{R}}.$$

考虑到 R 的定义, 只需证明, 对于任意自然数 $x, y, x', y', u, v, u', v'$, 关系

$$x+v = y+u \quad \text{和} \quad x'+v' = y'+u'$$

蕴含

$$x+x'+v+v' = y+y'+u+u'$$

和

$$xx'+yy'+uv'+u'v = xy'+x'y+uu'+vv'.$$

这无疑是容易的. 和与乘积已经在集合 $\mathbf{Z}$ 内定义, 我们必须证实这些运算的基本性质, 即

(I) 对于任意 $x, y, z \in \mathbf{Z}, x+y = y+x, x+(y+z) = (x+y)+z$; 存在 $\mathbf{Z}$ 的唯一的一个元素 0, 使得对于任意 $x \in \mathbf{Z}$ 有 $x+0 = x$.

(II) 对于任意 $x, y \in \mathbf{Z}$, 存在 $z \in \mathbf{Z}$ 使得 $x+z = y$.

(III) 对于任意 $x, y, z \in \mathbf{Z}, xy = yx, x(yz) = (xy)z$; 存在 $\mathbf{Z}$ 的唯一的一个元素 1, 使得对于任意 $x \in \mathbf{Z}$ 有 $1x = x$.

(IV) 对于任意 $x, y, z \in \mathbf{Z}, x(y+z) = xy+xz$.

例如, 我们证明 (IV). 在 $\mathbf{N} \times \mathbf{N}$ 里选择序偶使得

$$x = \mathrm{D}(x', x''), \quad y = \mathrm{D}(y', y''), \quad z = \mathrm{D}(z', z'');$$

于是有

$$\begin{aligned} x(y+z) &= \mathrm{D}(x', x'') \cdot \mathrm{D}(y'+z', y''+z'') \\ &= \mathrm{D}[x'(y'+z') + x''(y''+z''), x'(y''+z'') + x''(y'+z')] \end{aligned}$$

和

$$\begin{aligned} xy + xz &= \mathrm{D}(x'y' + x''y'', x'y'' + x''y') + \mathrm{D}(x'z' + x''z'', x'z'' + x''z') \\ &= \mathrm{D}(x'y' + x''y'' + x'z' + x''z'', x'y'' + x''y' + x'z'' + x''z'), \end{aligned}$$

比较所得的结果就得到 (IV).

最后还得把每个自然数 n 等同一个有理整数, 即

$$\mathrm{D}(n, 0).$$

容易验证这样定义的从 $\mathbf{N}$ 到 $\mathbf{Z}$ 内的映射 $n \to \mathrm{D}(n, 0)$ 是单射, 并且一方定义在自然数上的和另一方定义在有理整数上的代数运算是相容的. 这样一来, 把 $\mathbf{N}$ 看作 $\mathbf{Z}$ 的子集没有任何不便.

做了这件事, 我们注意到对于任意自然数 x 和 y 有

$$\mathrm{D}(x, y) = \mathrm{D}(x, 0) - \mathrm{D}(y, 0)$$

(有理整数 a 和 b 之间的差 $a - b$ 根据定义是唯一满足 $a = b + c$ 的有理整数 c: 见上述的性质 (II)). 事实上, 我们有

$$\mathrm{D}(x, y) + \mathrm{D}(y, 0) = \mathrm{D}(x + y, y),$$

所以问题归结为验证 $\mathrm{D}(x + y, y) = \mathrm{D}(x, 0)$, 即

$$(x + y) + 0 = y + x,$$

而这是显然的. 此后规定

$$\mathrm{D}(x, 0) = x \quad 对于所有 \ x \in \mathbf{N}$$

是方便的. 这样前面的关系就可以写作

$$\mathrm{D}(x, y) = x - y,$$

这就表明所有的有理整数是两个自然数的差.

设 z 是一个有理整数, 有理整数 $0 - z$ 记作

$$-z,$$

并且称为 z 的**相反数**, 其特征是

$$z + (-z) = 0,$$

显然有

$$\text{如果 } z = x - y, \text{ 则 } -z = y - x.$$

这些交代之后, 我们就写成 $z = x - y$, 其中 $x, y \in \mathbf{N}$. 有两种可能的情形:

一种是 $x \geqslant y$. 那么存在一个 $z' \in \mathbf{N}$ 使得 $x = y + z'$, 而由于这个关系在 $\mathbf{Z}$ 内也是成立的, 由于 $\mathbf{Z}$ 内减法的唯一性, 我们看到这时 $z = z'$, 换言之 z 是自然数.

另一种是 $y \geqslant x$. 令 $-z = y - x$ 得 $-z \in \mathbf{N}$.

因此, 对于所有有理整数 z, 或者 $z \in \mathbf{N}$, 或者 $-z \in \mathbf{Z}$. 说有理整数 $z \in \mathbf{N}$ 是非负的, 说其余的有理整数是非正的. 如果 x 和 y 是两个有理整数, 当 $y - x$ 非负时, 记为

$$x \leqslant y,$$

这样, $z \in \mathbf{Z}$ 是非负的, 其特征是

$$z \geqslant 0.$$

容易证明在集合 $\mathbf{Z}$ 里, 关系 "$\leqslant$" 具有显然的性质, 即:

关系 $x \leqslant y$ 和 $y \leqslant z$ 蕴含 $x \leqslant z$;

关系 $x = y$ 等价于 $x \leqslant y$ 且 $y \leqslant x$;

对于任意 x 和 y 有或者 $x \leqslant y$ 或者 $y \leqslant x$;

关系 $x \leqslant y$ 等价于 $x + z \leqslant y + z$;

如果 $z > 0, x \leqslant y$ 等价于 $xz \leqslant yz$; 如果 $z < 0$, $x \leqslant y$ 等价于 $xz \geqslant yz$.

以后我们有时会用到我们所没有陈述过的 $\mathbf{Z}$ 的性质, 一旦这些性质不是 "显然的", 我们会证明它们. 此刻, 我们建议读者不要过分深究本小节的内容, 这些内容的唯一目的是验证已经熟悉的结果, 没有这些结果, 数学研究无从谈起. 当读者渴望加深对于本书后面所陈述的内容的理解时, 他会有兴趣在相关的习题中写出本小节所忽略的所有证明的细节.

9. 有理数

构造了自然数, 随后构造了有理整数, 还必须构造**有理数**, 即两个有理整数的商 p/q, 其中 $q \neq 0$. 这些有理数的集合记作

$$\mathbf{Q}.$$

我们在这里不叙述如何从 $\mathbf{Z}$ 出发构造 $\mathbf{Q}$, 这是因为它跟从 $\mathbf{N}$ 出发构造 $\mathbf{Z}$ 所使用的方法类似, 更由于从 $\mathbf{Z}$ 出发构造 $\mathbf{Q}$ 是一个更一般的过程的特殊情形, 其详情将在 §29 介绍, 在其他情形我们也需要该一般过程.

我们给读者提出建议, 或者是容易地相信有理数的 "显然的" 性质 (有理数在本书中仅作为例子出现), 或者在阅读 §6, §7 和 §8 (这几节对于阅读 §29 是必需的) 后, 直接阅读 §29. 不言而喻, 对于初学者, 如果他不会产生错觉, 认为除他能够直接证明以外的所有性质都是平凡的, 那么还是使用第一种方法为好.

§5 习题

1. (G. Cantor) 设 X 是一个集合, 而 f 是从 X 到 X 的子集的集合 $\mathscr{P}(\mathrm{X})$ 内的一个映射. 用 A 表示满足条件 $x \notin f(x)$ 的 $x \in \mathrm{X}$ 的集合. 指出不存在 $x \in \mathrm{X}$, 使得 $\mathrm{A} = f(x)$. 由此推出不存在任何从 X 到 $\mathscr{P}(\mathrm{X})$ 内的满射, 随之对于任何基数 x 有

$$x < 2^x.$$

¶2. (G. Cantor) 设 $(f_n)_{n\in\mathbf{N}}$ 是从 $\mathbf{N}$ 到 $\mathbf{N}$ 自身内的映射的一个序列, 令

$$f(n) = f_n(n) + 1 \quad 对于所有\ n \in \mathbf{N},$$

定义一个从 $\mathbf{N}$ 到 $\mathbf{N}$ 内的映射. 指出不存在任何 $p \in \mathbf{N}$, 使得 $f = f_p$. 由此推出从 $\mathbf{N}$ 到 $\mathbf{N}$ 内的所有映射的集合是不可数的.

¶3. 可数个可数集合的并集是可数的 (借鉴第 5 小节的例 1 所使用的方法).

4. (G. Cantor) 设 I 是介于 0 和 1 之间的实数的集合. 用无限小数展开表示每一个 $x \in \mathrm{I}$ (可以用无穷个数字 0 跟随). 设 $(x_n)_{n\in\mathbf{N}}$ 是 I 的元素的一个序列, 组成一个数 $x \in \mathrm{I}$ 如下: 如果 x_n 的第 n 位数字异于 1, 则令 x 的第 n 位数字为 1, 如果 x_n 的第 n 位数字是 1, 则令 x 的第 n 位数字为 2. 指出对于所有 n 有 $x \neq x_n$. 由此得到结论: 集合 I (所有实数的集合 $\mathbf{R}$ 更不待说) 是不可数的.

说明这个推理与上面习题 2 的推理是相似的.

¶¶5. 设 f 是从集合 X 到集合 Y 的一个子集 Y_1 上的双射, 而 g 是从集合 Y 到集合 X 的一个子集 X_1 上的双射. 我们打算证明 X 和 Y 是等势的 (Bernstein). 为此定义 X 的子集 A_n 和 Y 的子集 B_n 如下: 令

$$\mathrm{A}_0 = \mathrm{X} - \mathrm{X}_1, \quad \mathrm{B}_1 = f(\mathrm{A}_0), \quad \mathrm{A}_1 = g(\mathrm{B}_1), \quad \mathrm{B}_2 = f(\mathrm{A}_1), \quad \mathrm{A}_2 = g(\mathrm{B}_2), \quad \cdots,$$

再定义一个映射 $h : \mathrm{X} \to \mathrm{Y}$ 如下: 给定一个 $x \in \mathrm{X}$, 取

$$h(x) = f(x), \quad 如果\ x \in \bigcup_{n\in\mathbf{N}} \mathrm{A}_n,$$

如果不然 (从而 $x \in \mathrm{X}_1$), 则令

$$h(x) = g^{-1}(x).$$

证明 h 是从 X 到 Y 上的一个双射.

6. 设 E 是有 n 个元素的有限集合, 而 n 是一个自然数. 考虑从 E 到 $\mathbf{N}$ 内的映射 f 使得 h 个数 $f(x)(x \in \mathrm{E})$ 的和至多等于 n, 证明所考虑的映射 f 的数目等于

$$\binom{n+h}{h}.$$

7. 证明对于自然数 n, 二项式系数 $\binom{n}{p}$ 的和等于 2^n.

¶8. 证明关系

$$\binom{n}{0}\binom{n}{p} + \binom{n}{1}\binom{n-1}{p-1} + \binom{n}{2}\binom{n-2}{p-2} + \cdots + \binom{n}{p}\binom{n-p}{0} = 2^p\binom{n}{p}$$

[首先研究, 在一个包含 n 个元素的集合 X 内, 有多少个包含 p 个元素的子集, 它含有预先指定的 k 个元素].

¶9. 设 p 和 n 是满足条件 $1 \leqslant p \leqslant n$ 的整数, 并且用 $S_{n,p}$ 表示从集合 $\{1, 2, \cdots, n\}$ 到集合 $\{1, 2, \cdots, p\}$ 的满射的数目. 证明

$$p^n = S_{n,p} + \binom{p}{1} S_{n,p-1} + \binom{p}{2} S_{n,p-2} + \cdots + \binom{p}{p-1}.$$

由此推出

$$S_{n,p} = p^n - \binom{p}{1}(p-1)^n + \binom{p}{2}(p-2)^n - \cdots + (-1)^{p-1}\binom{p}{p-1}.$$

对于 $p = 2, 3$ 化简这个结果.

¶¶¶10. 设 E 和 F 是两个有限集, 而 $x \to A(x)$ 是从 E 到 F 的子集的集合的一个映射. 证明存在从 E 到 F 内的满足条件

$$f(x) \in A(x), \quad 对于所有\ x \in E$$

的一个映射 f, 必须并且只需对于 E 的所有子集 H 有

$$\text{Card}\left(\bigcup_{x \in H} A(x)\right) \geqslant \text{Card}(H).$$

(这个结果一般以婚姻引理被熟知.)

11. 利用自然数的所有 (有限或无限) 非空集合存在最小整数这一事实证明下列经典结果:

a) 所有 $n \geqslant 2$ 的整数至少具有一个素因数 (我们提醒称一个整数 $p \geqslant 2$ 为**素数**, 如果它只有正因数 1 和 p).

b) 给定两个自然数 a 和 b, 其中的 $b \geqslant 1$, 则存在自然数 q 和 r, 使得

$$a = bq + r, \quad r < b,$$

并且整数 q 和 r 是唯一的 (**Euclid 除法**, 或 a 除以 b 的带余除法).

12. 证明素数集合是无限的.

¶13. 形如 $4n - 1$ (对应的, $6n - 1$) 的素数集合是无限的.

[这个结果是 Dirichlet 的算术数列定理的一个特殊情形, 这个定理说, 如果 a 和 b 是互素的整数, 则存在形如 $an + b$ 的无穷个素数. 这个数论中最著名的定理之一的 Dirichlet 定理的一般证明, 在目前还不能用纯算术的方法得到, 所有已经知道的证明 (包括 Dirichlet 本人的证明, 本质上不能简化) 都使用分析的方法, 或直接受分析启发而得到的方法.]

14. (底为 q 的计数法) 选择一个自然数 $q \geqslant 2$.

a) 设 x 是一个非零自然数. 指出存在一个并且仅一个整数 $n \geqslant 0$, 使得

$$q^n \leqslant x < q^{n+1},$$

再证明存在一个并且仅一个整数 a_n, 它满足条件

$$0 \leqslant a_n \leqslant q - 1, \quad a_n q^n \leqslant x < (a_n + 1)q^n.$$

b) 证明对于所有自然数 x, 存在一个并且仅一个整数序列 $a_0, a_1, \cdots, a_r, \cdots$, 满足下列条件:

1) 对于所有 $r \geqslant 0$ 有 $0 \leqslant a_r < q-1$;

2) 使得 $a_r \neq 0$ 的整数 r 的个数是有限的;

3) $x = a_0 + a_1 q + a_2 q^2 + \cdots + a_r q^r + \cdots$

[这个和看似有无穷多项, 其实由于加在 a_r 上的条件 2), 它除有限项以外为零.] 证明问题 a) 中的 n 是使得 $a_n \neq 0$ 的最大整数, 并且问题 a) 定义的整数 a_n 等于问题 b) 中的数 a_n. 序列

$$a_n a_{n-1} \cdots a_0$$

(这里涉及的不是乘积!) 称为 x **的在底为 q 的计数系统中的展开**.

c) 求出数 718 的 2 进计数系统里的展开 (即底 $q=2$).

d) 设 x 是未必是整数的正有理数. 证明存在整数序列

$$\cdots, a_n, a_{n-1}, \cdots, a_0, a_{-1}, a_{-2}, \cdots$$

满足下列条件:

1) 对于所有 $r \in \mathbf{Z}$ 有 $0 \leqslant a_r \leqslant q-1$.

2) 使得 $a_r \neq 0$ 的正整数 r 的个数是有限的.

3) 对于所有 $r \in \mathbf{Z}$, 我们有

$$a_r q^r + a_{r+1} q^{r+1} + \cdots \leqslant x \leqslant q^r + a_r q^r + a_{r+1} q^{r+1} + \cdots,$$

我们说

$$a_n a_{n-1} \cdots a_0 \cdot a_{-1} a_{-2} \cdots a_{-r} \cdots$$

(其中 n 是使得当 $r > n$ 时有 $a_r = 0$ 的最小的自然数 n) 是 x 的在底为 q 的计数系统中的展开.

e) 为了满足前面问题的条件 1) 和 2) 的整数 a_n 的一个族 $a_n (n \in \mathbf{Z})$ 是一个有理数在底为 q 的计数系统中的展开, 必须并且只需存在一个有理整数 r 和一个自然数 $k \geqslant 1$ (有理数的展开周期), 使得有

$$a_{n-k} = a_n, \quad \text{对于所有 } n \leqslant r.$$

15. 荣誉军团的统帅、纯理论物理金融开发有限责任公司的主席, 即总经理, 因纯和平目的的空间研究的发展而获取了各种利益, 在下诺曼底以每公顷 8000 法郎的价钱购买了 200 公顷的区域. 受这个例子的启发, 一个在该公司工作的水暖工每月挣 800 法郎, 决定将每年工资的十分之一投资到年利率 4% 的国库券. 他应当工作多少年才能够在下诺曼底获得 200 公顷的一个区域, 以便在那里平静地度过他的余生? (计算复利, 但是忽略可能的通货膨胀.)

16. 根据 1954 年 7 月 21 日的世界报, 印度支那战争的临时开支由下表提供 (单位是百万老法郎):

1946 : 101.8	1949 : 177.3	1952 : 427.6
1947 : 131.3	1950 : 258.3	1953 : 403.5
1948 : 136.3	1951 : 321	1954 : 428

用这个例子验证加法的基本性质 (结合律和交换律).

17. 在 1954 年 11 月, 在阿尔及利亚有 1230000 欧洲人和 8300000 本地人. 同时, 在阿尔及尔大学, 有 4548 名欧洲大学生, 557 名本地大学生. 以 1% 的误差计算欧洲人和本地人进入高等教育的比率.

第二章 群, 环, 域

众所周知, 代数的内容就是计算. 古典代数对于数做计算, 19 世纪和 20 世纪数学的发展则迫使数学家 (和物理学家) 越来越多地要对性质迥异的并且孤立于日常体验的数学对象进行计算. 这样人们就被引导到群、环和域的抽象概念. 群的概念出现在几何里 (变换群: 旋转, 平移, 相似变换, 等等), 在数论里 (Gauss, Dirichlet, Hermite 的 "二次型的可逆元" 的研究: 问题涉及的是给定一个整系数二次齐次多项式, 求所有使多项式保持不变的整系数的线性代换), 还出现在代数里 (代数方程理论中的 Galois 群), 在分析里 (Lie 群, 自同构函数), 还应该提到在物理学里 (Lorenz 群); 这说明当今群的概念的重要性不亚于集合和函数的概念. 至于域和环的概念, 这是在 19 世纪的代数数理论的研究引导数学家 (主要是 Dedekind) 形成的, 今天在代数和数论的所有分支里得到应用, 在分析的一些分支里也得到应用, 这些分支经常有可能采用合适的术语.

本部分最后有一节是关于 "复数" 的, 其在代数里的重要性确切地说是最小的, 反而是在分析中, 要是没有复数, 实际上什么也做不到. 在本书中引进复数的方法是用以构建代数中更一般的结构, 这是它的主要价值.

§6 运算

1. 运算, 结合性和交换性

给定一个集合 X, 所有从乘积集合 X × X 到集合 X 自身内的映射称为 X 上的**运算**. 直观地看来, X 上的运算在于令 X 的每个序偶 (x, y) 对应 X 的第三个元素, 它按照预先给定的规律依赖于 x 和 y.

在实际中, 为了表示运算使用 $(x,y)\to x+y$, 或 $(x,y)\to xy$, 或 $(x,y)\to x\wedge y$ 这类记号. 在本节我们经常使用记号 $\perp$, 目前在数学中其他地方从未用它表示特定运算 (因此适合用以表示不论什么运算).

设 $(x,y)\to x\perp y$ 是 X 上的一个运算. 我们说这个运算是**结合的**, 如果

$$\text{对于任何 } x,y,z\in \mathrm{X}, \text{ 有 } x\perp(y\perp z)=(x\perp y)\perp z.$$

在这种情形, 给定 X 的任意个元素 $x_1,x_2,\cdots,x_n$, 我们 (对于 n 归纳) 定义

$$x_1\perp x_2\perp\cdots\perp x_n=(x_1\perp x_2\perp\cdots\perp x_{n-1})\perp x_n,$$

这时, 对于所有满足 $1\leqslant p\leqslant n$ 的整数 p 有关系

$$x_1\perp x_2\perp\cdots\perp x_n=(x_1\perp\cdots\perp x_p)\perp(x_{p+1}\perp\cdots\perp x_n).$$

如果把所考虑运算像乘法一样记作 $(x,y)\to xy$, 我们就说使用**乘法记号**. 对于所有 $x\in \mathrm{X}$ 和所有整数 $n\geqslant 1$, 用公式

$$x^n=x\cdots x\ (n \text{ 个因子})$$

定义 x 的 n **次幂**, 这时对于任何整数 $p,q\geqslant 1$ 有

$$x^px^q=x^{p+q}.$$

反之, 如果把运算记作 $(x,y)\to x+y$, 我们就说使用**加法记号**, 对于所有 $x\in \mathrm{X}$ 和所有整数 $n\geqslant 1$, 定义

$$nx=x+\cdots+x\ (n \text{ 项}).$$

不言而喻, 这个记号所表示的概念与 n 次幂概念除记号不同外, 没有差别. 说了这些, 再回到上面提到的用乘法记号的幂的乘法公式, 用加法记号现在翻译成

$$px+qx=(p+q)x.$$

我们回到在一个集合 X 上的运算 $(x,y)\to x\perp y$. 我们说一个这样的运算是**交换的**, 如果

$$x\perp y=y\perp x\quad \text{对于任意的 } x,y\in \mathrm{X}.$$

在实际中, 乘法记号对于交换运算和非交换运算都适合, 加法记号仅用于交换的运算.

设 $(x,y)\to x\perp y$ 是集合 X 的一个结合的和交换的运算, 而 $(x_i)_{i\in \mathrm{I}}$ 是 X 的元素的有限族. 设 I 的元素个数是 n, 把这些元素写成一个序列的形式 $i_1,\cdots,i_n$, X 的元素

$$x_{i_1}\perp x_{i_2}\perp\cdots\perp x_{i_n}$$

(由于运算的结合性和交换性) 不依赖将 I 的元素写成 n 项序列的方式. 作为定义,令

$$\mathop{\perp}_{i\in \mathrm{I}} x_i = x_{i_1}\perp x_{i_2}\perp\cdots\perp x_{i_n}.$$

如果采用乘法记号, 则写成

$$\prod_{i\in \mathrm{I}} x_i = x_{i_1}x_{i_2}\cdots x_{i_n};$$

如果采用加法记号, 则利用记号

$$\sum_{i\in \mathrm{I}} x_i = x_{i_1}+x_{i_2}+\cdots+x_{i_n}.$$

注 1　刚引进的缩写记号在实际中经常被修改, 而这种修改用到时一看就会明白. 比如, 如果 I 是整数 $\{1,\cdots,n\}$ 组成的集合, 经常写成

$$x_1+\cdots+x_n=\sum_{i=1}^{n} x_i \quad 或 \quad \sum_{1\leqslant i\leqslant n} x_i;$$

如果 I 是整数序偶 (i,j) 的集合, 其中 $1\leqslant i\leqslant p, 1\leqslant j\leqslant q$. 如果用 x_{ij} 记所考虑的族的 "一般" 项, 经常使用记号

$$\sum_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant q}} x_{ij} \quad 代替 \quad \sum_{(i,j)\in \mathrm{I}} x_{ij}.$$

我们注意, 在这种情形, 运算的交换性写成

$$\sum_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant q}} x_{ij} = \sum_{1\leqslant i\leqslant p}\left(\sum_{1\leqslant j\leqslant q} x_{ij}\right),$$

右端表示和

$$(x_{11}+x_{12}+\cdots+x_{1q})+\cdots+(x_{p1}+x_{p2}+\cdots+x_{pq}).$$

这类缩写记号的使用对于回避冗长的公式是必需的.

最后注意结果中记号

$$\sum_{i\in \mathrm{I}} x_i$$

中字母 i 不起作用, 也不出现在结果中, 它只是指出要施行的一个运算 (即对于当 i 在 I 里变动时所得到的所有 x_i 求和), 可以用未曾使用过的 (这种谨慎对于避免大错是至关重要的) 其他字母代替它.

再回到 X 上一个运算 $(x,y)\to x\perp y$. 称所有元素 $e\in \mathrm{X}$ 为对于这个运算的**中性元**, 如果

$$x\perp e = e\perp x = x \quad 对于所有\ x\in \mathrm{X}.$$

定理 1 如果一个运算有一个中性元, 则它是唯一的.

事实上, 假定 e' 和 e'' 是中性元, 公式 $e' \perp x = x$ 特别给出 $e' \perp e'' = e''$; 公式 $x \perp e'' = x$ 特别给出 $e' \perp e'' = e'$; 故有 $e' = e''$, 定理得证.

现在给出运算的几个重要例子.

例 1 在有理整数 (即任意符号的整数) 集 $\mathbf{Z}$ 上, 我们有人人皆知的三个运算: 加法 $(x, y) \to x+y$, 它是交换的. 结合的, 并且有一个中性元 (即数 0); 乘法 $(x, y) \to xy$, 它是交换的. 结合的, 并且有一个中性元 (即数 1); 最后减法 $(x, y) \to x - y$, 它既不是交换的, 也不是结合的, 并且没有中性元.

在这个例子中可以用有理数集 $\mathbf{Q}$ 或实数集 $\mathbf{R}$ 代替 $\mathbf{Z}$.

例 2 设 $\mathbf{Q}^*$ 是非零有理数集, 在 $\mathbf{Q}^*$ 上的乘法 $(x, y) \to xy$ 是一个运算 (结合的、交换的, 并且有中性元); 除法 $(x, y) \to x/y$ 也是运算 (它既不是交换的, 也不是结合的, 并且没有中性元). 我们要注意除法不是所有有理数集合 $\mathbf{Q}$ 上的运算, 因为商 x/y 当 $y = 0$ 时没有定义, 除法不是定义在整个 $\mathbf{Q} \times \mathbf{Q}$ 上.

例 3 取 $\mathrm{X}=\mathbf{N}$, 这是自然数集合, 取映射 $(x, y) \to \mathrm{lcm}(x, y)$($x$ 和 y 的最小公倍数) 和映射 $(x, y) \to \gcd(x, y)$(x 和 y 的最大公约数). 这些是交换的和结合的运算, 第一个有中性元, 而第二个则没有中性元 (当然读者应当作为习题证明这些断言).

例 4 设 E 是任意一个集合, 而 X 则是从 E 到 E 内的所有映射的集合. 从 $\mathrm{X} \times \mathrm{X}$ 到 X 内的映射 $(f, g) \to f \circ g$ 是 X 上的一个运算, 它是结合的 (§2, 定理 2), 有一个中性元 (恒等映射 j_{X}), 但不是交换的.

例 5 设 E 是任意一个集合, 而 $\mathrm{X} = \mathscr{P}(\mathrm{E})$, 这是 E 的子集的集合, 那么映射 $(x, y) \to x \cap y$ 和 $(x, y) \to x \cup y$ 是 X 上的运算, 根据 §3, 第 1 小节的公式, 这些运算是结合的和交换的. E 是第一个运算的中性元, 而空集是第二个运算的中性元.

例 6 设 X 是空间 (通常意义下的空间) 里的起点为 O 的向量的集合, 令起点为 O 的所有向量 x, y 序偶对应它们的 "向量积" (随作者的不同, 记作 $x \times y$ 或 $x \wedge y$, 我们这里用 $x \wedge y$). 这样我们就得到一个运算, 它既不是结合的也不是交换的.

2. 可对称元

设 $(x, y) \to x \perp y$ 是集合 X 上的一个运算, 它有一个中性元 e. 给定一个元素 $x \in \mathrm{X}$, 称所有使得

$$x' \perp x = e \text{ (对应的, } x \perp x' = e)$$

成立的元素 $x' \in \mathrm{X}$ 为 x 的**左对称元** (对应的, **右对称元**), 称满足

$$x' \perp x = x \perp x' = e$$

的所有元素 $x' \in \mathrm{X}$ 为 x 的**对称元**. 如果存在 x 的对称元, 则说 x 是**可对称的**.

如果跟用乘法符号书写的运算打交道, 则用**逆元**这个词代替对称元, 用**可逆的**这个词代替可对称的 [例如, 如果考虑 $\mathbf{Q}$ 上的乘法 $(x,y) \to xy$, 可逆元就是非零有理数, 而一个这样的数的逆元则是 $1/x$]. X 的可逆元 x 的逆元一般记作

$$x^{-1}.$$

如果是跟采用加法记号的运算打交道, 则说**相反元**, 以代替对称元, 并且把 $x \in \mathrm{X}$ 的相反元记作

$$-x.$$

(无论如何, 当用 0 表示 X 的中性元时, 差不多总是使用加法记号的情形.)

最后要注意对称元. 左对称元和右对称元仅在非交换运算的情形有区别.

我们给出一个例子, 它指出在非交换情形, 这三个概念是有区别的.

例 7 取例 4 的运算, 说一个元素 $f \in \mathrm{X}$ 有一个*左逆元*, 意指存在一个映射 $g \in \mathrm{X}$ 使得 $g \circ f = j_{\mathrm{E}}$; 为此, 必须且只需 f 是单射的 (§2, 定理 3). 说一个元素 $f \in \mathrm{X}$ 有一个*右逆元*, 意指存在一个映射 $g \in \mathrm{X}$ 使得 $f \circ g = j_{\mathrm{E}}$. 为此, 必须且只需 f 是满射的 (§2, 定理 4). 最后说 f 是*可逆的*, 显然意指 f 是双射的, 而对于所考虑的运算, f 的*逆元*就是 §2 的意义下的逆映射.

定理 2 *设 $(x,y) \to x \perp y$ 是集合 E 上的一个结合的, 并且有一个中性元的运算. 为了使 E 的一个元素 x 是可对称的, 必须且只需它有一个左对称元和一个右对称元; 此时 x 有一个唯一的对称元, 它也是 x 的唯一的左对称元和唯一的右对称元.*

设 x' 是 x 的一个左对称元, 而 x'' 是 x 的一个右对称元, 于是有

$$x' \perp x = x \perp x'' = e,$$

这里 e 是 E 的中性元. 考虑到结合性, 我们从上面的等式推出

$$x'' = e \perp x'' = (x' \perp x) \perp x'' = x' \perp (x \perp x'') = x' \perp e = x'.$$

因此, x 的每一个左对称元等于它的每一个右对称元, 这就表明 x 有唯一的左对称元和唯一的右对称元, 并且它们是相等的, 用 x' 表示它们的公共值, 则有

$$x' \perp x = x \perp x' = e.$$

所以 x 是可对称的, 并且以 x' 为其对称元 (必然是唯一的, 因为对称元更是左对称元和右对称元). 这就完成了证明.

从此处至本节末尾, 我们假定在集合 E 上的运算是结合的并且有一个中性元 e. 对于 E 的一个可对称元 x, 可以谈论它的对称元 x' (或在乘法记号时, 逆元 x^{-1}; 在加法记号时, 相反元 $-x$).

定理 3　如果 $x \in \mathrm{E}$ 是可对称的, 则其对称元 x' 也是对称的, 而 x' 的对称元就是 x. 如果 x 和 y 是可对称的, 则 $x \perp y$ 也是可对称的, 并且有

$$(x \perp y)' = y' \perp x'.$$

关系

$$x' \perp x = x \perp x' = e$$

使得第一个断言是平凡的. 为了证明第二个断言, 我们做计算

$$(y' \perp x') \perp (x \perp y) = y' \perp (x' \perp x) \perp y = y' \perp e \perp y = y' \perp y = e,$$
$$(x \perp y) \perp (y' \perp x') = x \perp (y \perp y') \perp x' = x \perp e \perp x' = x \perp x' = e.$$

这就证明了 $x \perp y$ 是可对称的, 并且其对称元恰是 $y' \perp x'$.

使用乘法记号, 定理 3 翻译如下: 如果 x 是可逆的, 则其逆元 x^{-1} 也是可逆的, 并且

$$(x^{-1})^{-1} = x;$$

如果 x 和 y 是可逆的, 则 xy 也是可逆的, 并且有

$$(xy)^{-1} = y^{-1}x^{-1}.$$

使用加法记号, 我们另外假设运算是交换的. 这时定理 3 翻译如下: 如果 x 有相反元, 则其相反元 $-x$ 也有相反元, 并且

$$-(-x) = x;$$

如果 x 和 y 都有相反元, 则 $x+y$ 也有相反元, 并且有

$$-(x+y) = (-x) + (-y),$$

一般把此式写成

$$-(x+y) = -x - y.$$

定理 4　设 a 是 E 的一个可对称元, 那么对于 $b \in \mathrm{E}$, 存在唯一的一个 $x \in \mathrm{E}$ 使得

$$a \perp x = b,$$

即

$$x = a' \perp b.$$

事实上, 关系 $a \perp x = b$ 蕴含 $a' \perp (a \perp x) = a' \perp b$, 即

$$a' \perp b = (a' \perp a) \perp x = e \perp x = x.$$

反之, 从 $x = a'\perp b$ 得到

$$a\perp x = a\perp(a'\perp b) = (a\perp a')\perp b = e\perp b = b.$$

这就完成了证明.

使用乘法记号: 如果 a 是可逆的, 则方程

$$ax = b$$

有且仅有一个解, 即

$$x = a^{-1}b;$$

使用加法记号: 如果 a 有相反元, 则方程

$$a + x = b$$

有且仅有一个解, 即

$$x = b + (-a),$$

还可以写成

$$x = b - a$$

(b 和 a 之间的**差**).

作为本节的结束, 我们指出在本书后面, 不再使用符号 $\perp$、"可对称的" 和 "对称元" 等这些词. 我们总是跟记成乘法的和加法的运算打交道, 对于这样的运算再使用 "可对称的" 和 "对称元" (比如像初学者那样, 对于乘法谈论一个非零实数的对称元) 这些词将是十分滑稽可笑的.

§7 群的概念

1. 群的定义, 例子

称由一个集合 G 和 G 上的一个运算 $(x, y) \to xy$ 组成的序偶为群, 如果该运算满足下列条件:

a) 对于任意 $x, y, z \in \mathrm{G}, x(yz) = (xy)z$ (结合性);

b) 存在 G 的一个元素 e, 使得对于所有 $x \in \mathrm{G}$ 有 $xe = ex = x$ (中性元的存在性);

c) 对于所有 $x \in \mathrm{G}$ 存在一个元素 $x^{-1} \in \mathrm{G}$, 使得 $x^{-1}x = xx^{-1} = e$ (对于 G 的所有元, 逆元的存在性).

为了定义一个群, 只给出集合 G 是不够的, 还必须给出 G 上的一个运算, 并且验证上述的条件 a), b), c). 不过我们总用表示集合的同样字母, 比如 G 表示群, 其实 G 仅是组成给定条件的一部分.

初学者还要留意群并**不是** "一个其上存在一个满足条件 a), b), c) 的运算的集合", 因为我们可以容易地证明: 在所有集合上存在一个这样的运算, 甚至可以构造无穷多个这类运算, 只要给定的集合本身是无限集. 因此说 "群是一个其上存在一个这样的运算的集合", 就相当于说 "一个群就是一个集合", 不言而喻, 这样愚蠢透顶.

事实上, 在群的理论中, 人们对于在一个给定的集合 G 上一个满足条件 a), b), c) 的运算的存在性不感兴趣, 反之, 人们假定了这样的运算一劳永逸地预先给定, 并且打算利用它证明定理.

在上述定义中, 我们采用了乘法记号, 事实上人们确实常常这样做 (我们过去指出中性元有时代替 e 记作 1, 并且经常称为**单位元**). 但是当我们有一个**交换群** (或称 **Abel 群**) 时, 即一个运算是交换的群, 有时使用加法记号 $(x, y) \to x + y$; 在这种情形, 条件 a), b), c) 翻译如下:

a′) 对于任意 $x, y, z \in \mathrm{G}$, $x + (y + z) = (x + y) + z$;

b′) 存在 G 的一个元素 0, 使得对于所有 $x \in \mathrm{G}$ 有 $x + 0 = x$ (中性元的存在性);

c′) 对于所有 $x \in \mathrm{G}$ 存在一个元素 $-x$, 使得 $x + (-x) = 0$.

当然, 在这种情形, 必须加上条件

d′) 对于任意 $x, y \in \mathrm{G}$, 有 $x + y = y + x$.

例 1 有理整数集 **Z** 和加法 $(x, y) \to x + y$ 构成一个交换群, 称为**有理整数的加法群**. 用 **Q** 或 **R** 代替 **Z**, 同样定义**有理数的加法群**和**实数的加法群**.

例 2 由非零有理数的集合 $\mathbf{Q}^*$ 和这个集合上的运算 $(x, y) \to xy$ 形成的序偶是一个群 (其中性元是数 1), 称为**非零有理数的乘法群**. 同样定义**非零实数的乘法群** $\mathbf{R}^*$.

例 3 用 $\mathbf{Q}_+^*$ 表示由**正**有理数集合和在这个集合上的普通乘法运算得到的群. 同样用 $\mathbf{R}_+^*$ 表示正实数的乘法群.

注意, 反之, 使得 $0 < x \leqslant 1$ 的实数 x 的集合 I 和这个集合上的乘法 $(x, y) \to xy$ 形成的序偶不是一个群: 群的条件 c) 不满足.

例 4 设 X 是任意一个集合, 我们曾经谈到 (§2, 第 8 小节) 称从 X 到 X 内的所有双射为 X 的置换. 设 $\mathfrak{S}(\mathrm{X})$ 是这些置换的集合, 如果 f 和 g 是 X 的置换, 复合映射 $f \circ g$ 同样是 X 的置换 (§2, 定理 6), 于是公式 $(f, g) \to f \circ g$ 定义集合 $\mathfrak{S}(\mathrm{X})$ 上的一个运算. 这个运算是结合的 (§2, 定理 2); 它有一个中性元, 即恒等映射 j_{X} (经常称为集合 X 的单位置换); 最后, 如果 f 是 X 的一个置换, 则 (根据 §2, 定理 5) 逆映射 f^{-1} 也是一个置换, 而对于所考虑的运算它显然是逆元.

这样一来, 由集合 $\mathfrak{S}(\mathrm{X})$ 和这个集合上的运算 $(f, g) \to f \circ g$ 形成的序偶是一个群, 称为**集合X 的置换群**. 历史上正是对于这个群 (当 X 是一个有限集) 的研究 Galois 导出群的一般的和 "抽象的" 概念.

作为例子, 取 X 是由 1, 2, 3 组成的集合, 那么 $\mathfrak{S}(X)$ 含有 6 个元素, 即置换

$$
\begin{aligned}
s_1 &: 1,2,3 \to 1,2,3 \\
s_2 &: 1,2,3 \to 2,3,1 \\
s_3 &: 1,2,3 \to 3,1,2 \\
s_4 &: 1,2,3 \to 1,3,2 \\
s_5 &: 1,2,3 \to 2,1,3 \\
s_6 &: 1,2,3 \to 3,2,1
\end{aligned}
$$

而运算由下列 “乘法表” 给出:

	s_1	s_2	s_3	s_4	s_5	s_6
s_1	s_1	s_2	s_3	s_4	s_5	s_6
s_2	s_2	s_3	s_1	s_5	s_6	s_4
s_3	s_3	s_1	s_2	s_6	s_4	s_5
s_4	s_4	s_6	s_5	s_1	s_3	s_2
s_5	s_5	s_4	s_6	s_2	s_1	s_3
s_6	s_6	s_5	s_4	s_3	s_2	s_1

(我们采取下列约定: 表中的乘积 $s_i s_j$ 是第 i 行第 j 列的元素. 例如: $s_2 s_4 = s_5, s_4 s_2 = s_6$).

这个例子证实了**有限群**的存在性, 所谓有限群, 即其集合的元素个数为有限的群; 显然 $\mathfrak{S}(X)$ 是有限的当且仅当 X 是有限的.

当 X 是自然数 $1, 2, \cdots, n$ 组成的集合时 (上面看到 $n = 3$ 的情形), 代替记号 $\mathfrak{S}(X)$, 我们使用记号

$$\mathfrak{S}_n,$$

称 $\mathfrak{S}_n$ 为 n **个对象的置换群**或 n **个变量的对称群**. 我们在 §5 (定理 9 的推论) 看到 $\mathfrak{S}_n$ 的元素个数是整数

$$n! = 1 \cdot 2 \cdot \cdots \cdot n,$$

即前 n 个正整数的乘积. 我们有

$$6! = 720, \quad 7! = 5040, \quad 8! = 40\,320, \quad 9! = 362880, \quad 10! = 3628800,$$

希望通过观察 $\mathfrak{S}_n$ 的乘法表导出它的性质是天方夜谭.

例 5 考虑通常空间的给定起点 O 的向量的集合 G 的 (采用加法记号) 交换群, 在这个集合上由古典的平行四边形法则给定运算 $(x, y) \to x + y$.

2. 群的直积

设 $\mathrm{G}_1, \cdots, \mathrm{G}_n$ 是采取乘法记号的群. 在乘积集合 (§2, 第 2 小节)

$$\mathrm{G} = \mathrm{G}_1 \times \cdots \times \mathrm{G}_n$$

上考虑运算

$$(x_1, \cdots, x_n)(y_1, \cdots, y_n) = (x_1y_1, \cdots, x_ny_n),$$

由集合 G 和这个运算形成的序偶**是一个群**.

为了验证结合性, 考虑 G 的三个元素

$$x = (x_1, \cdots, x_n), \quad y = (y_1, \cdots, y_n), \quad z = (z_1, \cdots, z_n),$$

根据定义

$$(x_1, \cdots, x_n)(y_1, \cdots, y_n) = (x_1y_1, \cdots, x_ny_n), yz = (y_1z_1, \cdots, y_nz_n),$$

有

$$(xy)z = ((x_1y_1)z_1, \cdots, (x_ny_n)z_n), \quad x(yz) = (x_1(y_1z_1), \cdots, x_n(y_nz_n)),$$

因此, 从给定在 $\mathrm{G}_1, \cdots, \mathrm{G}_n$ 上的结合性得到在 G 内的结合性.

为了指出 G 具有一个中性元, 只需考虑元素

$$e = (e_1, \cdots, e_n),$$

其中对于 $1 \leqslant i \leqslant n$, e_i 表示 G_i 的中性元. 一个平凡的计算立刻指出, 对于 G 上的运算, e 是中性元. 最后, 如果

$$x = (x_1, \cdots, x_n)$$

是 G 的一个元素, 直接看出 x 有一个由公式

$$x^{-1} = (x_1^{-1}, \cdots, x_n^{-1})$$

给定的逆元. 因此我们通过赋予乘积集合 $\mathrm{G}_1 \times \cdots \times \mathrm{G}_n$ 以上面定义的运算得到一个群, 称这样得到的群为**群** $\mathrm{G}_1, \cdots, \mathrm{G}_n$ **的直积**.

当群 $\mathrm{G}_1, \cdots, \mathrm{G}_n$ 是交换的并且采用加法记号时, 我们对于它们的直积也采用加法记号 (这是合理的, 因为交换群的直积仍是交换的). 直积上的运算由

$$(x_1, \cdots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n)$$

给定, 并且直积的中性元是

$$(0, \cdots, 0)$$

(这里用同一个符号 0 表示不同群 $G_1, \cdots, G_n$ 的中性元).

最后, 给定一个群 G, 对于所有整数 $n \geqslant 1$, 群

$$G^n$$

是 n 个等于 G 的群的直积:

$$G^n = G \times \cdots \times G \ (n \text{ 个因子}).$$

例 6　加法群 $\mathbf{Z}^n$ 定义如下: 它的元素是 n 个有理整数的序列 $(x_1, \cdots, x_n)$, 而运算由

$$(x_1, \cdots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n)$$

给定. 同样定义**加法群 $\mathbf{Q}^n$** ($\mathbf{Q}$ 是在例 1 中定义的有理数的加法群) 和**加法群 $\mathbf{R}^n$** ($\mathbf{R}$ 是实数加法群).

在一个平面上, 选择两个坐标轴 Ox, Oy, 令 $\mathbf{R}^2$ 的每个元素 (x, y) 对应起点在 O 的向量 $\overrightarrow{OP}$, 相对于选定的坐标系, 它有坐标 (x, y). 这样我们就得到了从集合 $\mathbf{R}^2$ 到平面上的起点为 O 的向量的集合上的一个双射. 这个双射变换 $\mathbf{R}^2$ 的两个元素 (x', y') 和 (x'', y'') (在 $\mathbf{R}^2$ 的加法群里) 的和为以它们为坐标的向量 $\overrightarrow{OP'}$ 和 $\overrightarrow{OP''}$ 的和: 事实上, 众所周知为了向量相加, 应当把它们的分量相加.

于是, 我们可以把 $\mathbf{R}^2$ 内的加法看成是平面上两个向量的和的 “几何” 概念的代数翻译.

3. 群的子群

我们说 G 的一个子集 H 是 G 的一个**子群**, 如果 H 是非空的, 并且

$$x \in H \text{ 和 } y \in H \quad \text{蕴含} \quad xy^{-1} \in H.$$

一个群 G 至少总具有两个子群: G 本身和缩减为 G 的中性元的集合 $\{e\}$ (如果 G 由一个元素组成, 则这两个子群当然是重合的). 另外, 显然加法群 $\mathbf{Z}$ 和 $\mathbf{Q}$ 都是加法群 $\mathbf{R}$ 的子群 (例 1).

设 H 是群 G 的一个子群. 由于 H 是非空的, 它至少含有一个元素 a, 从关系 $a \in H$ 和 $a \in H$ 得到 H 含有

$$aa^{-1} = e;$$

这样, G 的一个子群 H 总含有 G 的中性元. 此外, 设 H 含有一个元素 x, 由于它含有 e 和 x, 它也含有

$$ex^{-1} = x^{-1},$$

这样, 关系 $x \in H$ 蕴含关系 $x^{-1} \in H$. 设 x, y 是 H 的元素, 根据前面提到的, H 具有元素 x, y^{-1}, 它就应当具有

$$x(y^{-1})^{-1} = xy,$$

从而关系 $x \in \mathrm{H}$ 和 $y \in \mathrm{H}$ 蕴含 $xy \in \mathrm{H}$.

反之, 考虑 G 的一个子集 H, 如果它具有下列三个性质:

a) 关系 $x \in \mathrm{H}$ 和 $y \in \mathrm{H}$ 蕴含 $xy \in \mathrm{H}$;

b) H 含有 G 的中性元 e;

c) 关系 $x \in \mathrm{H}$ 蕴含关系 $x^{-1} \in \mathrm{H}$.

那么 H 是 G 的一个子群. 事实上, 根据 b), H 是非空的; 此外, 设 x, y 是 H 的两个元素, 根据 c), H 含有 x 和 y^{-1}, 根据 a), 它含有 xy^{-1}; 我们的断言得以验证.

如此看来, 上面的条件 a), b) 和 c) 表征了子群, 在实际中经常用它们代替原来的定义.

反之应当注意, G 的一个子集 H 仅满足 a), 或仅满足 a) 和 b), 未必是 G 的一个子群. 比如 G 是有理整数的加法群 **Z**, $n \geqslant 0$ 的整数集合 **N** 满足 a) 和 b), 但不是 **Z** 的一个子群.

设 H 是 G 的一个子群, 上面的条件 a) 指出从 $\mathrm{G} \times \mathrm{G}$ 到 G 内的映射 $(x, y) \rightarrow xy$ 映射 $\mathrm{H} \times \mathrm{H}$ 到 H 内, 从而 "诱导" 出 H 上的一个运算. 交代了这些, 集合 H 配备了这个运算之后成为一个群. 事实上, 给定在 G 上的运算是结合的, 在 H 上也是结合的; 因为 H 具有 G 的中性元, 那么看作 H 上的运算, 显然也具有一个中性元 (即 G 的中性元); 最后, 由于上面的条件 c), 所有 $x \in \mathrm{H}$ 在 H 内是可逆的.

从此以后, 当我们考虑一个群 G 的一个子群 H 时, 总是把 H 看作一个像上述那样配备了由 G 的运算诱导的运算的群.

例 7 给定一个任意集合 X, X 的置换群 $\mathfrak{S}(\mathrm{X})$ 的所有子群称为**集合 X 的变换群**. 因而 X 的一个变换群是从 X 到 X 内的映射的集合 G, 这些映射具有下列性质: 所有 $s \in \mathrm{G}$ 是双射的; G 含有恒等映射 j_{X}; 并且如果 G 含有两个映射 s 和 t, 则也含有 st^{-1}. 于是可以把 G 看作一个群, 其中配备运算

$$(s, t) \rightarrow s \circ t.$$

初等几何提供了许多变换群的例子: 直线上、平面上或空间里的平移群; 绕平面上或空间里一个固定点的旋转群; 平面上或空间里的位移群; 平面上或空间里的中心给定和非零比例的位似变换群; 等等.

例 8 取整数加法群 **Z** 作为 G. **Z** 的一个子群是整数的一个集合 I, 它满足下列条件: $0 \in \mathrm{I}$, 如果 I 含有两个整数 x 和 y, 则也含有 $x - y$. 对于所有整数 n, n 的倍数的集合记作 $n\mathbf{Z}$ (即整数 nx 的集合, x 遍历 **Z**), 显然这是 **Z** 的一个子群. 反之, 对于 **Z** 的所有子群 I, 存在唯一的一个整数 $n \geqslant 0$ 使得 $\mathrm{I} = n\mathbf{Z}$. 首先如果 $\mathrm{I} = \{0\}$, 这个断言是平凡的: 只需取 $n = 0$. 其次假定 $\mathrm{I} \neq \{0\}$, 在 I 内存在非零整数, 实际上是正整数 (因为如果 $n \in \mathrm{I}$, 则 $-n \in \mathrm{I}$). 设 n 是属于 I 的最小的正整数 (我们回忆, 根据 §5, 注 2, 在非负整数的所有的集合里存在一个最小的元素), 我们要证明 $\mathrm{I} = n\mathbf{Z}$.

事实上, 由于 I 含有 n, 它含有 $n+n=2n$, 于是 $n+2n=3n$, 等等, 从而对于所有 $x \geqslant 1$, 它含有 nx; 此外 I 含有 $n0=0$; 最后, 如果 x 是一个负整数, 根据已经证明的, I 含有 $n(-x)=-nx$, 于是也含有 $-(-nx)=nx$; 这样我们证明了包含关系 $n\mathbf{Z} \subset \mathrm{I}$. 剩下需要证明的是反向的包含. 为此考虑一个元素 $x \in \mathrm{I}$, 并且写出 (Euclid 除法)

$$x = nq + r \quad (0 \leqslant r < n).$$

I 含有 n, 故含有 nq, 由于它含有 x, 故含有 $x-nq=r$; r 是正的或 0, 并且小于 n; 如果 $r \neq 0$, n 将不是 I 中的正整数中的最小者, 故必有 $r=0$, $x=nq$, 这就证明 I 的所有元素都是 n 的一个倍数, 换句话说, $\mathrm{I} \subset n\mathbf{Z}$, 故必有 $\mathrm{I}=n\mathbf{Z}$.

为了证明 n 的唯一性, 只需证明如果 $p\mathbf{Z}=q\mathbf{Z}$, 其中 $p \geqslant 0, q \geqslant 0$, 则有 $p=q$. 由于 $q\mathbf{Z}$ 含有 q, 所做的假设指出 q 是 p 的倍数, 同样 p 也是 q 的倍数, 由此显然得到 $p=q$.

$\mathbf{Z}$ 的所有子群都是 $n\mathbf{Z}$ 这种形式这个事实在数论和其他地方起着重要的作用, 并且在许多情形, 可以方便地代替基于整数的 "Euclid 除法" (或 "带余数除法") 所做的证明.

作为例子, 我们指出这个结果如何导出最大公约数 (h.c.f.) 的重要性质. 设 $x_1, \cdots, x_n$ 是非零整数, 用 I 表示这样的 $x \in \mathbf{Z}$ 的集合, 存在 $u_1, \cdots, u_n \in \mathbf{Z}$ 使得

$$x = u_1x_1 + \cdots + u_nx_n.$$

显然 I 是 $\mathbf{Z}$ 的一个子群, 且 $\mathrm{I} \neq \{0\}$ (因为每个 $x_i \in \mathrm{I}$), 因此 $\mathrm{I}=d\mathbf{Z}$, 这里 d 是一个完全确定的正整数. I 的所有元素是 d 的一个倍数; 一个特殊的推论是: $x_1, \cdots, x_n$ 是 d 的倍数, 故 d 是给定的整数的一个公约数. 但是另一方面, $x_1, \cdots, x_n$ 的所有公约数 d' 显然整除 $u_1x_1+\cdots+u_nx_n$, 其中 $u_1, \cdots, u_n$ 是任意整数, 因此整除 I 的所有元素, 特殊情形是整除 d. 换句话说, d 是 $x_1, \cdots, x_n$ 的最大公约数, 同时看出 d 可以写成形式

$$d = u_1x_1 + \cdots + u_nx_n,$$

其中的 u_i $(1 \leqslant i \leqslant n)$ 是适当选取的整数.

我们注意, 作为推论的 Bezout 定理: *$x_1, \cdots, x_n$ 是互素的, 必须且只需存在整数 $u_1, \cdots, u_n$, 满足*

$$u_1x_1 + \cdots + u_nx_n = 1.$$

事实上, 采用上面的记号, 这个条件表达的是子群 I 含有数 1. 如果这个条件满足, x_i 显然互素; 如果 x_i 互素, 则它们的最大公约数 $d=1$, 由于 $\mathrm{I}=d\mathbf{Z}$, 故 1 是 d 的倍数, 即 $d=1$.

考虑 $\mathbf{Z}$ 的子群

$$x_1\mathbf{Z} \cap \cdots \cap x_n\mathbf{Z}$$

同样得到关于 n 个数 $x_1, \cdots, x_n$ 的最小公倍数的性质. 如果用 m 表示这个子群的正的生成元, 可以直接验证 m 就是给定整数 x_i 的最小公倍数.

这个问题在 §31 还会以更仔细更一般的方式进行研究.

例 9 $\mathbf{Z}$ 的子群 $n\mathbf{Z}$ 的构造推广如下: 设 G 是一个群 (采用乘法记号, 因为下面不需要交换性), 而 x 是 G 的一个元素. 对于所有有理整数 p, 如下定义 x^p:

$$x^p = \begin{cases} x \cdots x \ (p \text{ 个因子}), & \text{如果 } p \geqslant 1, \\ e \ (\text{中性元}), & \text{如果 } p = 0, \\ (x^{-1})^{-p}, & \text{如果 } p < 0. \end{cases}$$

借助 G 内的乘法的结合性, 容易验证下列计算法则:

$$x^p x^q = x^{p+q}, \quad (x^p)^{-1} = x^{-p}, \quad (x^p)^q = x^{pq}.$$

由此看出 x^p 的集合 (x 给定, 而 p 在 $\mathbf{Z}$ 内变化) 是 G 的一个子群: 事实上, 它显然不是空集, 又如果它含有元素 $u = x^p, v = x^q$, 公式 $uv^{-1} = x^{p-q}$ 指出它含有 uv^{-1}.

这个子群称为 G 的**由 x 生成的子群** (因此在加法群 $\mathbf{Z}$ 里, $n\mathbf{Z}$ 是由 n 生成的子群), 其元素称为 x 的**幂**.

当 G 采用加法记号时, 用记号 px 代替记号 x^p, 并且说 px 是 x 的**整倍元**. 根据定义

$$px = \begin{cases} x + \cdots + x \ (p \text{ 项}), & \text{如果 } p \geqslant 1, \\ 0 \ (\text{中性元}), & \text{如果 } p = 0, \\ (-p)(-x), & \text{如果 } p < 0. \end{cases}$$

并且有公式

$$px + qx = (p+q)x, \quad -(px) = (-p)x, \quad p(qx) = (pq)x.$$



注 1 在加法群内还有关系: 对于所考虑的群的任意元素 x 和 y, 所有的有理整数 p, 有

$$px + py = p(x+y).$$

在任意群 (因此用乘法记号) 中, 类似的公式

$$x^p y^p = (xy)^p$$

是错误的, 除非 x 和 y 是**交换的**, 即如果有 $xy = yx$, 则上式成立. 首先, 如果 $xy = yx$, 则

$$(xy)^2 = xyxy = xxyy = x^2y^2,$$

然后如此继续下去. 反之, 如果当 $p=2$ 时 $(xy)^p = x^p y^p$ 成立, 那么 $xyxy = xxyy$, 以 x^{-1} 左乘和以 y^{-1} 右乘等式两端, 即得 $x^{-1}xyxyy^{-1} = x^{-1}xxyyy^{-1}$, 此式可改写为 $xy = yx$, 这正是预料到的.

注 2 如果存在一个 $x \in \mathrm{G}$, 使得群 G 的所有元素都是 x 的幂, 则称这个群是*循环群*, 并且说 x 是 G 的一个*生成元*. 加法群 **Z** 是循环的, 并且有生成元 1 和 -1. 存在有限的循环群 (考虑任意一个有限群 G 和由 G 的任意一个元素生成的子群), 后面将会看到可以完全描述它们的结构.

4. 子群的交, 生成元

我们有下列结果:

定理 1 *设 $(\mathrm{H}_i)_{i\in \mathrm{I}}$ 是群 G 的子群的一个族, 那么 H_i 的交集还是 G 的一个子群. 为了使 H_i 的并集还是 G 的一个子群, 只需对于任意指标 $i, j \in \mathrm{I}$, 存在一个指标 $k \in \mathrm{I}$ 使得*

$$\mathrm{H}_i, \mathrm{H}_j \subset \mathrm{H}_k.$$

设 M 是 H_i 的交集, 那么 M 是非空的 (因为 G 的中性元属于每个 H_i, 从而属于 M). 如果 M 含有两个元素 x 和 y, 那么它们属于每个 H_i, xy^{-1} 也是这样, 从而 xy^{-1} 也属于 M, 于是 M 是 G 的一个子群.

现在设 U 是 H_i 的并集, 它显然不是空集. 设 x 和 y 是 U 的两个元素, 那么存在指标 $i, j \in \mathrm{I}$ 使得 $x \in \mathrm{H}_i, y \in \mathrm{H}_j$. 根据定理中的假设, 存在一个指标 $k \in \mathrm{I}$ 使得 H_k 同时包含 x 和 y, 因此包含 xy^{-1}, 从而 xy^{-1} 属于并集 U, 这就完成了证明.

设 B 是 G 的一个子集, 存在 G 的一个子群包含 B (比如, G 自己). 根据定理 1 所有这些子群的交集还是 G 的一个子群, 并且还包含 B, 由构造本身知道它包含在 G 的所有包含 B 的子群内. 这个交子群因而是 G 的包含 B 的子群中的 "最小者", 我们说这是 G 的**由 B 生成的子群**.

假设 B 缩减为仅有一个元素 x, 含有 x 的一个子群显然含有 x 的所有的幂 (在上面的例 9 中定义). 而这些幂组成了 G 的一个含有 x 的从而包含 B 的子群. 于是我们这里看到 G 的包含 B 的最小子群就是由 x 的幂组成的子群, 即在例 9 的意义下由 x 生成的子群.

在 G 的任意子集 B 的情形下, 可以采用类似例 9 的方法构造由 B 生成的子群:

定理 2 *设 B 是群 G 的一个子集. $x \in \mathrm{G}$ 属于 G 的由 B 生成的子群, 必须且只需存在一个整数 $p \geqslant 0$ 和元素 $x_1, \cdots, x_p \in \mathrm{G}$ 使得*

a) 有关系

$$x = x_1 \cdots x_p.$$

b) 对于每个 i $(1 \leqslant i \leqslant p)$, 有 $x_i \in \mathrm{B}$ 或 $x_i^{-1} \in \mathrm{B}$.



注 3　对于 $p=0$, 作为约定, 定理陈述的条件 a) 应当解释为 $x=e$ [一般, 我们约定在一个群里空乘积 (或零个因子的乘积) 就是群的中性元. 这只是约定, 而为了保证某些陈述的有效性这个约定是必需的].

为了证明定理 2, 考虑满足定理中所陈述的条件的 $x\in \mathrm{G}$ 的集合 H: 所有的事情归结为证明 H 是一个子群, 包含 B, 并且包含于所有包含 B 的子群内.

第三个断言是显然的: 如果一个子群包含 B, 它显然含有断言 b) 的元素, 从而含有断言 a) 中所表示的元素 x.

H 包含 B 这件事是明显的: 事实上 B 的元素 x 满足条件 a) 和 b), 只需取 $p=1$ 和 $x_1=x$.

剩下的是证明 H 是一个子群. 首先根据上面的注 3, H 含有中性元. 再设 x 和 y 是 H 的两个元素, 于是可以写出

$$x=x_1\cdots x_p,\quad y=y_1\cdots y_q,$$

满足条件

$$\text{对于每个 } 1\leqslant i\leqslant p,\quad x_i\in \mathrm{B} \text{ 或 } x_i^{-1}\in \mathrm{B};$$
$$\text{对于每个 } 1\leqslant j\leqslant q,\quad y_j\in \mathrm{B} \text{ 或 } y_j^{-1}\in \mathrm{B}.$$

于是根据 §6 中的定理 3,

$$xy^{-1}=(x_1\cdots x_p)(y_1\cdots y_q)^{-1}=x_1\cdots x_p y_q^{-1}\cdots y_1^{-1},$$

这样我们就把 G 的元素 xy^{-1} 分解成乘积

$$xy^{-1}=z_1\cdots z_{p+q},$$

其中

$$z_k\in \mathrm{B} \text{ 或 } z_k^{-1}\in \mathrm{B},\quad \text{当 } 1\leqslant k\leqslant p+q.$$

这就证明了 $xy^{-1}\in \mathrm{H}$. 因此, H 是 G 的一个子群, 定理证毕.

当 G 的由一个子集 B 生成的子群是整个 G 时, 就说 B 是 G 的**一个生成元集**. 如果 G 拥有一个生成元的有限集 (即存在 G 的一个生成 G 的有限子集 B), 则说 G 是一个**有限生成群** —— 显然所有循环群是有限生成的.

设 G 是一个有限生成的交换群, 而 $\mathrm{B}=\{a_1,\cdots,a_n\}$ 是 G 的有限生成元集. 利用定理 2 得到: 所有 $x\in \mathrm{G}$ 有一个分解

$$x=x_1\cdots x_p,\ \text{对于每个 } i \text{ 有 } x_i\in \mathrm{B} \text{ 或 } x_i^{-1}\in \mathrm{B};$$

这个分解中的每个因子或是 a_j 中的一个, 或是 a_j^{-1} 中的一个. 而由于 G 是可交换的, 对于一个给定的指标 j, 我们可以把等于 a_j 或 a_j^{-1} 的所有 x_i 放在一起, 这些 x_i

的乘积显然是 a_j 的一个幂, 最终我们得到 x 的一个形式为

$$x = a_1^{r_1} \cdots a_n^{r_n}$$

的分解, 其中 r_i 是有理整数.

反之, 如果所有的 $x \in \mathrm{G}$ 可以写成上述形式, 那么显然 G 是有限生成的并且是由元素 $a_1, \cdots, a_n$ 生成的.

例 10 加法群 $\mathbf{Z}^n$ 是有限生成的, 并且这个群的生成元集合的元素是

$$e_1 = (1, 0, \cdots, 0), \quad e_2 = (0, 1, \cdots, 0), \quad \cdots, \quad e_n = (0, \cdots, 0, 1).$$

事实上, 如果 $r_1, r_2, \cdots, r_n$ 是任意有理整数, 直接有公式

$$\begin{aligned} r_1 e_1 &= (r_1, 0, \cdots, 0), \\ r_2 e_2 &= (0, r_2, \cdots, 0), \\ &\cdots\cdots\cdots\cdots \\ r_n e_n &= (0, 0, \cdots, r_n), \end{aligned}$$

因此有

$$r_1 e_1 + r_2 e_2 + \cdots + r_n e_n = (r_1, r_2, \cdots, r_n),$$

这就表明 $\mathbf{Z}^n$ 的所有元素都是元素 $e_1, e_2, \cdots, e_n$ 的幂 (这里是倍数) 的乘积 (这里是和).

反之, 有理数的加法群 $\mathbf{Q}$ 不是有限生成的. 事实上, 如果 $\mathbf{Q}$ 是由有限个有理数

$$a_1 = p_1/q_1, \cdots, a_n = p_n/q_m$$

生成的, 这就是说, 对于所有的有理数 x, 都存在有理整数 $r_1, r_2, \cdots, r_n$ 使得

$$x = r_1 a_1 + \cdots + r_n a_n.$$

由此显然 x 可以写成一个有公分母 $q_1 \cdots q_n$ (或更一般的 $a_1, \cdots, a_n$ 的任何公分母) 的分数形式. 换句话说, 将可能把所有有理数同时化成有同样的分母的形式, 而这看起来是荒谬的(*)!

5. 置换和对换

考虑集合

$$\mathrm{I}_n = \{1, 2, \cdots, n\}$$

(*) 但是还是建议读者证明这一断言.

的置换群 $\mathfrak{S}_n$, 我们说置换 $t \in \mathfrak{S}_n$ 是一个对换(*), 如果存在满足 $1 \leqslant i \leqslant n-1$ 的一个整数 i, 使得有下列关系:

$$t(i) = i+1, \quad t(i+1) = i, \quad \text{对于 } k \neq i, i+1 \text{ 有 } t(k) = k.$$

定理 3　群 $\mathfrak{S}_n$ 是由其含有的对换生成的.

事实上, 我们要用对于 n 的归纳法推理证明所有置换 $s \in \mathfrak{S}_n$ 是对换的一个乘积. $n=1$ 的情形是平凡的, 因为这时 $\mathfrak{S}_1$ 缩减为只含有一个中性元.

考虑一个置换 $s \in \mathfrak{S}_n$, 并且令 $s(n) = i$. 用 t_j 表示把 j 变成 $j+1$ 的对换, 显然置换

$$u = t_{n-1} \circ \cdots \circ t_i \circ s$$

满足 $u(n) = n$, 并且由于对于所有的对换有

$$t^{-1} = t,$$

故我们有

$$s = t_i^{-1} \circ \cdots \circ t_{n-1}^{-1} \circ u = t_i \circ \cdots \circ t_{n-1} \circ u.$$

为了证明 s 是对换的一个乘积, 只需对于 u 证明这一结论, 而 u 是一个满足 $u(n) = n$ 的置换. 而最后这个关系指出 u 置换 I_n 的 $1, 2, \cdots, n-1$, 换句话说, u 在 I_{n-1} 中 "诱导" 出一个置换 $u' \in \mathfrak{S}_{n-1}$. 根据归纳假设, u' 可以写成

$$u' = v_1 \circ \cdots \circ v_q,$$

其中的 $v_1, \cdots, v_q$ 是群 $\mathfrak{S}_{n-1}$ 内的对换. 定义 I_n 的置换 $w_1, \cdots, w_q$ 为

$$w_j(x) = \begin{cases} v_j(x), & \text{如果 } x \in \mathrm{I}_{n-1}, \\ n, & \text{如果 } x = n. \end{cases}$$

由于 u 和 u' 在 I_{n-1} 内重合, 显然有

$$u = w_1 \circ \cdots \circ w_q.$$

由于 v_j 是 I_{n-1} 的对换, 故显然 w_j 是 I_n 的对换. 从而在群 $\mathfrak{S}_n$ 中, 置换 u 是对换的乘积, 这就完成了证明.

6. 陪集

设 G 是一个群, 而 H 是 G 的一个子群, 那么关系

$$\mathrm{R}\{x, y\} : x^{-1}y \in \mathrm{H}$$

(*) 这里的对换指的只是一类特殊的对换. —— 译者注

是 §4, 第 1 小节意义下的集合 G 上的一个等价关系. 首先, 显然 R$\{x,x\}$ 总是真的, 因为它表示 H 含有 G 的中性元. 其次, 为了证明 R$\{x,y\}$ 蕴含 R$\{y,x\}$, 根据子群定义, 我们注意到

$$x^{-1}y \in \mathrm{H} \text{ 蕴含 } (x^{-1}y)^{-1} \in \mathrm{H}, \text{ 即 } y^{-1}x \in \mathrm{H}.$$

最后, 关系 R$\{x,y\}$ 和 R$\{y,z\}$ 就是关系

$$x^{-1}y \in \mathrm{H} \quad \text{和} \quad y^{-1}z \in \mathrm{H},$$

由子群的定义得到关系

$$(x^{-1}y)(y^{-1}z) \in \mathrm{H}, \quad \text{即} \quad x^{-1}z \in \mathrm{H},$$

即关系 R$\{x,z\}$.

从而所考虑的关系 R$\{x,y\}$ 确实是 G 上的一个等价关系. 我们要构造对应的等价类 F_x(§4, 第 2 小节). 对于 $x \in \mathrm{G}$, 根据定义, 集合 F_x 是使关系 R$\{x,y\}$ 为真的 $y \in \mathrm{G}$ 组成的集合, 换句话说, 有关系 $x^{-1}y \in \mathrm{H}$. 令 $x^{-1}y = z$, 则有 $y = xz$, 这就是说关系 R$\{x,y\}$ 为真的意义就是 $z \in \mathrm{H}$. F_x 是形式为 xz 的 G 的元素的集合, 其中 $z \in \mathrm{H}$. 基于这个理由, 我们用记号

$$x\mathrm{H}$$

代替记号 F_x, 并且称为 H 的左陪集 (同样定义 H 的右陪集, 这是形如 Hx 的 G 的子集, 这里 Hx 表示形如 zx 的元素的集合, 其中的 $z \in \mathrm{H}$). H 的左陪集 xH (对应的, 右陪集 Hx) 的集合, 即集合 G 关于等价关系 $x^{-1}y \in \mathrm{H}$ (对应的, $yx^{-1} \in \mathrm{H}$) 的商集记作 G/H (对应的 H \ G).

例 11 取有理整数加法群 $\mathbf{Z}$ 作为 G, 取 $p\mathbf{Z}$ 作为 H, $p\mathbf{Z}$ 是由一个给定的整数 p 的倍数组成的子群. 那么关系 R$\{x,y\}$ 表示为 (采用加法记号)

$$y - x \in p\mathbf{Z} \quad \text{即} \quad x \equiv y \pmod{p},$$

这就回到 §4 的例 4, 这里子群 $p\mathbf{Z}$ 的陪集正是 §4 的例 9 定义的模 p 的同余类.

关于这个主题, 我们注意在 §4 的例 4 中曾经在 $\mathbf{Z}/p\mathbf{Z}$ 上定义一个 "加法", 这个运算满足关系: 对于任意 $x, y \in \mathbf{Z}$, 有

$$\theta(x+y) = \theta(x) + \theta(y),$$

(θ 表示从 $\mathbf{Z}$ 到 $\mathbf{Z}/p\mathbf{Z}$ 上的典范映射). 可以容易地证明 (见 §8, 第 3 小节) $\mathbf{Z}/p\mathbf{Z}$ 配备了这个运算构成一个群 (模 p 的整数加法群). 在本节的习题 16 里可以看到一个更一般的构造方法.

我们注意到, 对于所有 H 的左陪集 xH, 存在从 H 到 xH 上的一个*双射*, 即映射 $z \to xz$ (根据类的定义这个映射是满射的; 由于所有 $x \in \mathrm{G}$ 是可逆的, 应用 §6 定理

4 得到单射性). 如果在特殊情形下, G 是一个有限群, 这时 H 也是有限的, 我们看到每个左陪集 xH 跟 H 含有同样个数的元素. 而左陪集 xH 组成集合 G 的一个划分. 于是 (§5, 定理 7) G 的元素个数等于 H 的元素个数乘以不同的左陪集 xH 的个数. 因此, 有

定理 4 如果 G 是一个有限群, 而 H 是 G 的一个子群. 则有

$$\mathrm{Card}(\mathrm{G}) = \mathrm{Card}(\mathrm{G}/\mathrm{H}) \times \mathrm{Card}(\mathrm{H}).$$

注 4 有限群 G 的元素个数 Card(G) 称为 G **的阶** (于是 "五阶有限群" 就是具有五个元素的有限群). 另外, 数 Card(G/H) 称为 H **在** G **内的指标**. 容易证明这也是 G 内的不同右陪集 Hx 的个数.

定理 4 表明 H 的指标是 G 的阶的一个约数. 我们将给出这个结果的一个重要应用.

定理 5 设 G 是一个 n 阶有限群, 则对于所有的 $x \in \mathrm{G}$ 有

$$x^n = e.$$

事实上, 设 H 是由 x 生成的 G 的子群, 并且设 $r = \mathrm{Card(H)}$, 由于 n 是 r 的倍数, 只需证明 $x^r = e$. 换句话说, 只需在 G 是由 x 生成的情形下证明定理 5, 以下就假定是这种情形.

那么考虑由

$$f(q) = x^q$$

给定的映射 $f : \mathbf{Z} \to \mathrm{G}$. 由假设它是满射的. 幂的运算法则 (例 9) 表明我们有关系

$$f(0) = e, \quad f(q' - q'') = f(q')f(q'')^{-1}.$$

从这些关系得到使得 $f(q) = e$ 的 $q \in \mathbf{Z}$ 组成 $\mathbf{Z}$ 的一个子群, 故有形式 $s\mathbf{Z}$, 这里 s 是一个完全确定的正整数.

进而, 关系 $f(q') = f(q'')$ 等价于

$$f(q')f(q'')^{-1} = e,$$

根据前面的等式 $f(q' - q'') = f(q')f(q'')^{-1}$, 上式可以写成 $f(q' - q'') = e$, 这就等价于

$$q' - q'' \in s\mathbf{Z}, \quad 即 \quad q' \equiv q'' \pmod{s}.$$

由于 f 是满射的, 那么 G 与 $\mathbf{Z}$ 里模 s 的类的集合有同样个数的元素. 换句话说, s 就是 G 的元素个数, 由于我们有 $x^s = e$, 定理证毕.

7. n 个对象的置换数

我们在 §5 (定理 10 的推论) 已经确立了以下结果:

设 X 是一个 n 个元素的有限集, 则 X 的置换群 $\mathfrak{S}(\mathrm{X})$ 的阶是

$$n! = 1 \cdot 2 \cdots n.$$

我们要给这个结果一个证明, 这个证明与 §5 中的证明没有本质区别, 只不过更系统地利用集合 $\mathfrak{S}(\mathrm{X})$ 上的群结构.

如果 $n = 1$ 定理是显然成立的, 我们要用关于 n 的归纳推理, 换句话说, 证明如果它对于整数 $n-1$ 时为真, 则对于整数 n 也为真.

为此, 选定 X 的一个元素 a, 并且设

$$\mathrm{Y} = \mathrm{X} - \{a\}$$

是从 X 去掉 a 得到的集合, Y 是一个有 $n-1$ 个元素的集合, 因此对于它上述定理是可以应用的 (归纳假设).

另外, 可以把 $\mathfrak{S}(\mathrm{Y})$ 看作 $\mathfrak{S}(\mathrm{X})$ 的一个子群. 为此只需令集合 Y 的所有置换 s 对应 X 的置换 $\bar{s}$, 其定义是

$$\bar{s} = \begin{cases} s(x), & \text{如果 } x \in \mathrm{Y}, \\ a, & \text{如果 } x = a; \end{cases}$$

按照这种方式, $\mathfrak{S}(\mathrm{Y})$ 等同于 $\mathfrak{S}(\mathrm{X})$ 的使得 a 为不动点的置换组成的子群.

根据归纳假设, 群 $\mathfrak{S}(\mathrm{Y})$ 具有 $(n-1)!$ 个元素. 为了由此推出 $\mathfrak{S}(\mathrm{X})$ 具有 $n!$ 个元素, 即前者的 n 倍, (根据定理 4) 只需证明在 $\mathfrak{S}(\mathrm{X})$ 里 $\mathfrak{S}(\mathrm{Y})$ 的左陪集正好有 n 个.

为此, 引进映射 $f : \mathfrak{S}(\mathrm{X}) \to \mathrm{X}$, 其定义是

$$f(s) = s(a) \quad \text{对于所有 } s \in \mathfrak{S}(\mathrm{X}).$$

给定 X 的置换 s 和 t, 关系 $f(s) = f(t)$ 写成

$$s(a) = t(a), \quad \text{即} \quad a = s^{-1}t(a),$$

这就是说

$$s^{-1}t \in \mathfrak{S}(\mathrm{Y}),$$

即 s 和 t 属于子群 $\mathfrak{S}(\mathrm{Y})$ 的同一个左陪集. 利用 §4 的定理 2 我们看到 $\mathfrak{S}(\mathrm{Y})$ 的左陪集的个数等于 $\mathfrak{S}(\mathrm{X})$ 在 f 下的像的个数, 即这样的元素 $x \in \mathrm{X}$ 的个数, 对于这些 x, 存在 X 的置换 s 使得 $x = s(a)$. 但是显然所有的 $x \in \mathrm{X}$ 都可以选取适当的置换 s 写成 $x = s(a)$, 因此, 在 $\mathfrak{S}(\mathrm{X})$ 内 $\mathfrak{S}(\mathrm{Y})$ 的左陪集的个数是 n, 这就结束了定理的证明.

8. 群的同态

给定群 G 和 H, 如果从 G 到 H 内的映射 f 满足条件

$$f(xy) = f(x)f(y) \quad \text{对于任意 } x, y \in \mathrm{G},$$

则称 f 为一个**从 G 到 H 内的同态**. 在这个等式中取 $y = e$, 即得 $f(x) = f(x)f(e)$, 因此有

$$f(e) = e.$$

再取 $y = x^{-1}$, 并且考虑到刚得到的结果, 即得

$$f(x^{-1}) = f(x)^{-1} \quad \text{对于所有 } x \in \mathrm{G}.$$

注 5 上面给出的定义假定了群 G 和 H 都是采用乘法记号, 如果 G 和 H 中有一个或两个采用加法记号, 则定义要适当修改. 例如, 如果 G 采用加法记号, 而 H 采用乘法记号, 一个从 G 到 H 内的映射 f 是一个同态, 如果满足

$$f(x + y) = f(x)f(y) \quad \text{对于任意 } x, y \in \mathrm{G}.$$

例 12 取有理整数的加法群 $\mathbf{Z}$ 作为 G, 取一个任意 (乘法) 群作为 H, 映射 f 由

$$f(n) = a^n \quad \text{对于任意 } n \in \mathbf{Z}$$

给定. 这是一个同态: 这个结论由例 9 的公式得到. 此外, 从 $\mathbf{Z}$ 到 H 内的所有同态 f 都用刚提到的方法给定. 事实上, 对于这样一个同态, 令

$$f(1) = a,$$

我们有

$$f(2) = f(1 + 1) = f(1)f(1) = aa = a^2,$$
$$f(3) = f(2 + 1) = f(2)f(1) = a^2a = a^3,$$

等等, 故对于正整数 n 有 $f(n) = a^n$, 对于负整数 n, 有

$$f(n) = f(-n)^{-1} = (a^{-n})^{-1} = a^n,$$

由此即得对于所有的 $n \in \mathbf{Z}$, 都有 $f(n) = a^n$.

例 13 对于所有整数 p, 从 $\mathbf{Z}$ 到 $\mathbf{Z}/p\mathbf{Z}$ 上的典范映射是从加法群 $\mathbf{Z}$ 到加法群 $\mathbf{Z}/p\mathbf{Z}$ 上的一个同态.

例 14 公式

$$\log(xy) = \log(x) + \log(y)$$

表明在分析中定义的对数函数是从乘法群 $\mathbf{R}_+^*$ 到加法群 $\mathbf{R}$ 的一个同态.

定理 6 设 $f: \mathrm{M} \to \mathrm{N}$ 和 $g: \mathrm{N} \to \mathrm{P}$ 是群的同态, 则 $g \circ f: \mathrm{M} \to \mathrm{P}$ 也是同态. 如果 $f: \mathrm{M} \to \mathrm{N}$ 是双射的, 则逆映射 $f^{-1}: \mathrm{N} \to \mathrm{M}$ 也是同态.

为了证明第一个断言, 只需注意到对于 $x, y \in \mathrm{M}$ 有

$$(g \circ f)(xy) = g(f(xy)) = g(f(x)f(y)) = g(f(x))g(f(y)).$$

为了证明第二个断言, 换句话说, 证明对于 $u, v \in \mathrm{N}$ 有

$$f^{-1}(uv) = f^{-1}(u)f^{-1}(v).$$

我们注意到只需证明两端在 f 下的像是相等的 (因为 f 是双射的, 当然更是单射的), 即证明

$$f(f^{-1}(uv)) = f(f^{-1}(u)f^{-1}(v)).$$

由 f^{-1} 的定义左端等于 uv; 而由于 f 是同态, 右端等于

$$f(f^{-1}(u))f(f^{-1}(v)) = uv,$$

这就结束了证明.

设 G 和 H 是两个群, 称从 G 到 H 上的所有双射的同态为**从 G 到 H 上的同构**, 并且说 **G 和 H 是同构的**, 如果存在从 G 到 H 上的同构.

例 15 利用例 14 中的对数函数 $x \to \log x$, 我们发现群 $\mathbf{R}_+^*$ 和 $\mathbf{R}$ 是同构的.

关系

G 和 H 是同构的

是一个等价关系. 事实上, 首先, 对于任意群 G, G 和 G 是同构的, 因为恒等映射显然是从 G 到 G 上的同构; 其次, 如果存在一个从 G 到 H 上的同构 f, 那么同样存在一个从 H 到 G 上的同构, 这就是 f^{-1}; 最后如果存在一个从群 M 到群 N 上的同构 f, 又存在一个从 N 到 P 上的同构 g, 那么存在一个从 M 到 P 上的同构, 这就是 $g \circ f$, 因为根据定理 6 它是一个同态, 由于 f 和 g 是双射, $g \circ f$ 也是双射.

称所有从 G 到 G 上的同构为群 G 的**自同构**.

例 16 设 G 是一个采用乘法记号的群, 那么对所有 $a \in \mathrm{G}$, 由

$$f(x) = axa^{-1}$$

给定的从 G 到 G 内的映射是 G 的一个自同构.

事实上, 我们有

$$
\begin{aligned}
f(x)f(y) &= (axa^{-1})(aya^{-1}) = (ax)(a^{-1}a)(ya^{-1}) \\
&= (ax)(ya^{-1}) = a(xy)a^{-1} = f(xy),
\end{aligned}
$$

由此得到 f 是一个同态. 进而, 对于所有 $y \in \mathrm{G}$, 方程 $axa^{-1} = y$ 有且仅有一个解 $x = a^{-1}ya$, 这表明 f 是双射的.

刚描述的方法得到的 G 的自同构称为群 G 的**内自同构**. 仅对于非交换的群这个概念才是有意义的.

例 17 考虑正实数的乘法群 $\mathbf{R}_+^*$, 那么对于所有非零实数 α, 在分析中定义的函数

$$f(x) = x^{\alpha}$$

是群 $\mathbf{R}_+^*$ 的一个自同构; 反自同构是

$$f^{-1}(x) = x^{1/\alpha}.$$



注 6 在实际中, 经常认为两个同构的群 G 和 H 是相同的. 更准确地说, 在纯粹群论的观点内, G 和 H 具有完全相同的性质. 比如, 如果 G 是交换的, 则 H 亦然; 如果 G 是由 n 个元素生成的, 则 H 也是; 一般, 一旦选定了从 G 到 H 上的一个同构, 就可以把 G 的元素之间的所有关系 “翻译” 成为 H 的对应元素之间的类似关系.

我们注意到对数函数, 作为乘法群 $\mathbf{R}_+^*$ 到加法群 $\mathbf{R}$ 的同构, 正是为了把正实数之间的乘法关系变换为任意符号的实数之间的加法关系而发明的.

9. 同态的核与像

首先确立下列结果:

定理 7 *设 f 是从一个群 G 到一个群 H 内的同态. G 的所有子群在 f 下的像是 H 的一个子群. H 的所有子群在 f 下的逆像是 G 的一个子群.*

设 G' 是 G 的一个子群, 而 $\mathrm{H}' = f(\mathrm{G}')$ 是它的像. 如果 $u, v \in \mathrm{H}'$, 那么存在 $x, y \in \mathrm{G}'$ 使得 $u = f(x), v = f(y)$, 于是有

$$uv^{-1} = f(x)f(y)^{-1} = f(x)f(y^{-1}) = f(xy^{-1}),$$

由于 $xy^{-1} \in \mathrm{G}'$, 因此有 $uv^{-1} \in \mathrm{H}'$, 这就证明了定理所陈述的第一个断言. 第二个断言的推理是类似的, 证明细节留给读者去完成.

从定理 7 得到, 如果 f 是从一个群 G 到一个群 H 内的同态, 那么 $f(\mathrm{G})$ 是 H 的一个子群, 称为 **f 的像**, 并且记作

$$\mathrm{Im}(f).$$

同样, 集合 $f^{-1}(\{e\})$ 由使得

$$f(x) = e$$

的 $x \in \mathrm{G}$ 组成, 它是 G 的一个子群, 称为 f **的核**, 记作

$$\mathrm{Ker}(f).$$

例 18 设 G 是一个乘法群, a 是 G 的一个元素, 考虑由 $f(n) = a^n$ 给定的同态 $f : \mathbf{Z} \to \mathrm{G}$. 那么 f 的像是 G 的由 a 生成的子群, 而 f 的核是 $\mathbf{Z}$ 的由满足 $a^n = e$ 的有理整数 n 组成的子群 (这些有理整数组成一个子群这个事实已经在定理 5 的证明中证实并使用).

注 7 设 N 是一个同态 $f : \mathrm{G} \to \mathrm{H}$ 的核. 对于 $x \in \mathrm{N}$ 和 $a \in \mathrm{G}$ 有



$$f(axa^{-1}) = f(a)f(x)f(a)^{-1} = f(a)ef(a)^{-1} = f(a)f(a)^{-1} = e,$$

因此有

$$axa^{-1} \in \mathrm{N} \quad 对于任意\ a \in \mathrm{G}\ 和\ x \in \mathrm{N}.$$

称群 G 的一个子群是**不变的** (或**正规的**), 如果它具有前述的性质, 这显然意味着对于 G 的所有内自同构 s 有

$$s(\mathrm{N}) \subset \mathrm{N}.$$

我们看到**一个同态的核是一个不变子群**. 反之, 可以证明, 如果 N 是群 G 的一个不变子群, 则存在一个群 H 和从 G 到 H 内的一个同态 f, 使得 N 是 f 的核. (见习题 16).

当 G 是交换的, 不言而喻 G 的所有子群都是不变的.

定理 8 设 G 和 H 是两个群, 而 f 是从 G 到 H 内的一个同态. 那么 f 是单射的, 必须且只需它的核只含有中性元.

由于 $f(e) = e$, 关系 $f(x) = e$ 即 $f(x) = f(e)$. 如果 f 是单射的, 这就蕴含 $x = e$, 换句话说, $\mathrm{Ker}(f) = \{e\}$. 反之, 假定 f 的核只含有 e, 关系 $f(x) = f(y)$ 还可以写成

$$f(x)f(y)^{-1} = e,$$

由于 f 是同态, $f(xy^{-1}) = e$, 这就意味着

$$xy^{-1} \in \mathrm{Ker}(f);$$

由于 $\mathrm{Ker}(f) = \{e\}$, 故 $xy^{-1} = e$, 即 $x = y$, 从而 f 是单射的, 这就完成了证明.

定理 9 设 G, H 和 M 是三个群, $p : \mathrm{G} \to \mathrm{H}$ 和 $f : \mathrm{G} \to \mathrm{M}$ 是同态. 假定 p 是满射的, 则下列条件是等价的:

a) 存在一个同态 $f' : \mathrm{H} \to \mathrm{M}$, 使得 $f = f' \circ p$;

b) $\mathrm{Ker}(p) \subset \mathrm{Ker}(f)$.

如果这些条件满足, 则同态 f' 是唯一的; 当且仅当 $\mathrm{Ker}(p) = \mathrm{Ker}(f)$, 它是单射的; 当且仅当 f 是满射的, f' 是满射的.

首先探求在什么条件下, 存在一个从 H 到 M 内的同态 f' 使得 $f = f' \circ p$, 回答由 §2 的定理 1 给出: 归根结底是要检查关系 $p(x) = p(y)$ 蕴含关系 $f(x) = f(y)$. 由于 p 是一个同态, 前一个关系改写为

$$e = p(x)p(y)^{-1} = p(xy^{-1}),$$

换句话说, $xy^{-1} \in \mathrm{Ker}(p)$. 同样的道理, 第二个关系改写为 $xy^{-1} \in \mathrm{Ker}(f)$. 取 $y = e$ 就可以看出, 关系 $x \in \mathrm{Ker}(p)$ 应当蕴含关系 $x \in \mathrm{Ker}(f)$, 这就是说 $\mathrm{Ker}(p) \subset \mathrm{Ker}(f)$, 而这个条件对于 $p(x) = p(y)$ 蕴含 $f(x) = f(y)$ 显然是充分的.

这样一来, 条件 b) 等价于存在一个映射 $f' : \mathrm{H} \to \mathrm{M}$, 使得 $f = f' \circ p$. 这个映射必然是一个同态. 事实上, 设 $u, v \in \mathrm{H}$, 由于 p 是满射的, 可以写出 $u = p(x)$, $v = p(y)$, 其中 $x, y \in \mathrm{G}$, 那么有

$$\begin{aligned} f'(uv) &= f'(p(x)p(y)) = f'(p(xy)) = f(xy) = f(x)f(y) \\ &= f'(p(x))f'(p(y)) = f'(u)f'(v), \end{aligned}$$

这就证明了我们的断言.

条件 a) 和 b) 的等价性得以证明.

f' 的唯一性是显然的, 因为, p 既然是满射的, 那么

$$\text{关系 } f_1' \circ p = f_2' \circ p \text{ 蕴含 } f_1' = f_2'.$$

显然有

$$f'(\mathrm{H}) = f'(p(\mathrm{G})) = f(\mathrm{G}),$$

因此, 当且仅当 f 是满射的时 f' 是满射的.

最后探求 f' 的核, 它由使得 $f'(u) = e$ 的 $u \in \mathrm{H}$ 组成. 令 $u = p(x)$, 这还可以写为 $f(x) = e$, 换句话说 $x \in \mathrm{Ker}(f)$. 因此

$$\mathrm{Ker}(f') = p(\mathrm{Ker}(f)).$$

f' 是单射的, $p(\mathrm{Ker}(f)) = \{e\}$ 是必要且充分的 (定理 8), 换句话说 $\mathrm{Ker}(f) \subset \mathrm{Ker}(p)$, 而 $\mathrm{Ker}(p) \subset \mathrm{Ker}(f)$ 已经被证明了, 故

$$\mathrm{Ker}(f) = \mathrm{Ker}(p),$$

这就结束了证明.

10. 应用到循环群

设 G 是一个循环群, 而 x 是 G 的一个生成元, 于是 G 的所有元素是 x 的一个幂. 换句话说, 由

$$f(n) = x^n$$

给定的同态 $f : \mathbf{Z} \to \mathrm{G}$ 是满射的.

用 I 表示 f 的核, I 是 $\mathbf{Z}$ 的一个子群, 因此存在唯一的一个整数 $p \geqslant 0$, 使得

$$\mathrm{I} = p\mathbf{Z}$$

(参见例 8). 区分两种情形:

第一种情形, $p = 0$. 那么由定理 8, f 是单射的, 从而是双射的, 因此是从加法群 $\mathbf{Z}$ 到 G 上的一个同构.

第二种情形, $p \neq 0$. 考虑加法群 $\mathbf{Z}/p\mathbf{Z}$ (例 11) 和从 $\mathbf{Z}$ 到 $\mathbf{Z}/p\mathbf{Z}$ 上的典范映射 g. 这是一个同态, 它的核是 $p\mathbf{Z}$, 由此得到 $\mathrm{Ker}(g) = \mathrm{Ker}(f)$. 根据定理 9, 存在唯一的一个同态

$$f' : \mathbf{Z}/p\mathbf{Z} \to \mathrm{G}$$

使得 $f = f' \circ g$. [这意味着对于所有整数 n, x^n 是模 np 的类在 f' 下的像; f' 的存在性来自事实:

$$\text{关系 } m \equiv n (\mathrm{mod}\ p) \text{ 蕴含 } x^m = x^n,$$

因此 x^n 不是依赖整数 n, 而是依赖它的模 p 的类. 当然, 这里的推理只不过是在定理 9 的证明中所使用的推理在当前情形的翻版.] 由于 f 是满射的, f' 也是满射的; 又由于 $\mathrm{Ker}(f) = \mathrm{Ker}(g)$, f' 是单射的, 于是 G 同构于整数模 p 的加法群 $\mathbf{Z}/p\mathbf{Z}$. 特别的, G 有与 $\mathbf{Z}/p\mathbf{Z}$ 同样的元素数, 而 $\mathbf{Z}/p\mathbf{Z}$ 有 p 个元素, 故 p 是 G 的元素个数. 故得

定理 10 *所有的无限循环群同构于加法群 $\mathbf{Z}$, 而所有有限循环群 G 同构于加法群 $\mathbf{Z}/p\mathbf{Z}$, 这里 p 是 G 的元素个数.*

由此显然推出: 两个循环群同构, 当且仅当它们有同样的元素数 (有限或无限).

设 x 是任意一个群 G 的一个元素, 称 G 的由 x 生成的子群 H 的阶 (或基数) 为 x **的阶**. 由于 H 是 $\mathbf{Z}$ 在同态 $n \to x^n$ 下的像, 我们发现, 为了 x 的阶是有限的, 必要且充分的条件是存在非零整数 p, 使得

$$x^p = e.$$

x 的阶是满足上述关系的最小整数 $p \geqslant 1$.

11. 作用在一个集合上的群

设 G 是一个群, 而 X 是一个集合. 我们称 G **作用在** X **上**, 如果给定从 $G \times X$ 到 X 内的一个映射, 记作

$$(s, x) \to s \cdot x,$$

满足两个条件: 结合关系

$$s \cdot (t \cdot x) = (st) \cdot x \quad 对于任意\ s, t \in G和\ x \in X,$$

和

$$e \cdot x = x \quad 对于任意\ x \in X,$$

这里 e 表示 G 的中性元.

例 19 可以让群 G 以多种方式作用在自身上, 或者借助映射 ("左平移")

$$(s, x) \to sx,$$

或者借助映射 ("右平移")

$$(s, x) = xs^{-1},$$

或者借助映射 ("内自同构")

$$(s, x) \to sxs^{-1}.$$

例 20 设 G 是一个群, 而 H 是 G 的一个子群, 取

$$X = G/H,$$

这是 G 内 H 的左陪集 xH 的集合 (第 6 小节). 对于 $s \in G$ 和 $A \in X$, 集合 $sA \subset G$ 仍然是 H 的一个左陪集, 这里 sA 是 sa 的集合, 其中 $a \in A$ (事实上, 如果取一个 $x \in A$, 那么 A 就是 xh 的集合, $h \in H$, 因此 sA 是 sxh 的集合. 换句话说, 如果 $A = sH$, 那么 $sA = (sx)H$, 这明显表明 $sA \in X$); 这就允许定义从 $G \times X$ 到 X 内的一个映射, 即 $(s, A) \to sA$. 从其构造本身可以直接验证, G 作用在 $X = G/H$ 上.

例 21 设 E 是一个集合, 而 p 是一个正整数; 令

$$X = E^p,$$

这是 E 的 p 个元素的序列 $(x_1, \cdots, x_p)$ 的集合, 而

$$G = \mathfrak{S}_p$$

是集合 $\{1, 2, \cdots, p\}$ 的置换群. 对于 $s \in G$, $x = (x_1, \cdots, x_p) \in X$ 定义

$$s \cdot x = (x_{s^{-1}(1)}, \cdots, x_{s^{-1}(p)}).$$

这样定义的从 G × X 到 X 内的映射允许 G 作用在 X 上. 事实上, 设 $s, t \in$ G, 而 $x \in$ X, 令

$$t \cdot x = y = (y_1, \cdots, y_p),$$

则有

$$s \cdot (t \cdot x) = s \cdot y = (y_{s^{-1}(1)}, \cdots, y_{s^{-1}(p)}).$$

但是

$$y = (x_{t^{-1}(1)}, \cdots, x_{t^{-1}(p)}),$$

故

$$y_i = x_{t^{-1}(i)}, \quad 1 \leqslant i \leqslant p;$$

因此有

$$s \cdot (t \cdot x) = (z_1, \cdots, z_p),$$

其中

$$z_i = y_{s^{-1}(i)} = x_{t^{-1}(s^{-1}(i))} = x_{(st)^{-1}(i)}.$$

这就证明了关系 $s \cdot (t \cdot x) = (st) \cdot x$; 关系 $e \cdot x = x$ 是显然的.

例 22 设 X 是一个集合, 而 G 是 X 的一个变换群 (例 7), 那么从 G × X 到 X 内的映射 $(s, x) \to s(x)$ 使得 G 作用在 X 上.

我们注意, 如果一个群 G 作用在一个集合 X 上, 那么对于所有 $s \in$ G, 由

$$\bar{s}(x) = s \cdot x$$

给定映射

$$\bar{s} : \mathrm{X} \to \mathrm{X},$$

由于 $s^{-1} \cdot (s \cdot x) = (s^{-1} \cdot s) \cdot x = e \cdot x = x$, 所以 $\bar{s}$ 是双射的. 这说明, 本小节开头所陈述的条件可以解释为: 映射 $s \to \bar{s}$ 是从群 G 到集合 X 的一个变换群上的同态.

设 G 是一个作用在集合 X 上的群. 对于每个 $x \in$ X, 使得 $s \cdot x = x$ 的 $s \in$ G 显然组成 G 的一个子群, 称它为 G**内**x 的**稳定子群**. 此外, 称形式为 $s \cdot x$ 的 X 的元素的集合为 x 的**轨道**, 其中的 $s \in$ G.

在本节的习题里会找到这些概念的补充内容.

§7 习题

1. 找出 一, 二或三个元素的所有的群.
2. 对于四个元素 (下面记作 e, a, b, c) 的集合配备由下列乘法表给定的具有交换性的运算:

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

证明这样就得到一个交换群. 求它的所有自同构 (这个群以 Klein **四元群**的名字被人熟知). 几何地解释这个群 (在空间内, 考虑关于三个直角构成的三面角的棱的对称).

3. 指出集合 $\{1, 2, 3, 4\}$ 的置换群 $\mathfrak{S}_4$ 具有一个同构于 Klein 四元群的不变子群.

4. 给实数集 $\mathbf{R}$ 配备运算

$$(x, y) \to \sqrt[3]{x^3 + y^3}.$$

指出这样得到一个同构于加法群 $\mathbf{R}$ 的群.

5. 设 $\mathrm{G}_1, \cdots, \mathrm{G}_n$ 是群, $\mathrm{H}_1, \cdots, \mathrm{H}_n$ 分别是 $\mathrm{G}_1, \cdots, \mathrm{G}_n$ 的子群; 指出 $\mathrm{H}_1 \times \cdots \times \mathrm{H}_n$ 是 $\mathrm{G}_1 \times \cdots \times \mathrm{G}_n$ 的子群.

6. 设 G 是以乘法标记的群. 对于所有 $a \in \mathrm{G}$, 令

$$s_a(x) = ax \quad 对于所有\ x \in \mathrm{G}$$

定义一个从 G 到 G 内的映射 s_a (在 G 内的*幅度为 a 的左平移*; 为了明白这个术语的起源, 读者可以考察这样一种情形: G 是通常空间的起点为给定的点 O 的所有向量的加法群). 指出映射 $a \to s_a$ 是从群 G 到集合 G 的置换群的一个同构.

7. 设 G 是一个 m 个元素的循环群, 而 x 是 G 的一个生成元. x^k 是 G 的一个生成元, 必须并且只需整数 m 和 k 是互素的 (利用 Bezout 定理). 在一般情形, G 的由 x^k 生成的子群的阶是多少?

8. 设 m 和 n 是两个有理整数. 存在一个整数 r, 使得有

$$r \equiv 0 (\mathrm{mod}\ m) \quad 和 \quad r \equiv 1 (\mathrm{mod}\ n),$$

必须并且只需 m 和 n 是互素的.

¶ 由此推出下列结果: 设 G 是一个交换群, x 和 y 分别是 G 的互素的 m 和 n 阶的元素, 那么 $z = xy$ 是 mn 阶的, 并且由 z 生成的 G 的子群含有 x 和 y. (称由 x 生成的子空间的元素数目为一个元素 x 的**阶**, 这个阶是有限的, 当且仅当存在一个整数 $n \neq 0$, 使得

$$x^n = e;$$

在这种情形, x 的阶是满足上式的最小的整数 $n \geqslant 1$, 读者可以证明这一事实.)

¶9. 设 G 和 H 分别是有 m 和 n 个元素的循环群. $\mathrm{G} \times \mathrm{H}$ 是循环群, 必须并且只需 m 和 n 是互素的. 在上述条件下, 如果 x 和 y 分别是 G 和 H 的生成元, 则序偶 (x, y) 是 $\mathrm{G} \times \mathrm{H}$ 的生成元.

¶10. 所有素数阶的有限群是循环的, 并且它的每个异于中性元的元素都是生成元 (利用 §7 的定理 4 或习题 7).

11. 设 A 是群 G 的一个子集. 称使得对于所有 $a \in \mathrm{A}$ 有 $xa = ax$ 的 $x \in \mathrm{G}$ 的集合 Z(A) 是 A 的**中心化子**. 指出 Z(A) 是 G 的一个子群. 指出 Z(G) (称为 G 的**中心**) 是 G 的一个交换和不变子群.

12. 一个群 G 的两个元素 x 和 y 称为**共轭的**, 如果存在一个 $s \in \mathrm{G}$, 使得

$$y = sxs^{-1}.$$

证明关系

$$x \text{ 和 } y \text{ 是共轭的}$$

是集合 G 上的一个等价关系. 取绕空间的一个给定点 O 的旋转群作为 G, 选择过 O 的一条直线 D, 指出 G 的所有的元素共轭于一个绕 D 的旋转.

13. 设 A 是群 G 的一个子集, 把对于一个 $x \in \mathrm{A}$ 形如 sxs^{-1} 的 G 的元素的集合记作 $s\mathrm{A}s^{-1}$ ($s \in \mathrm{G}$ 给定). 说明如果 A 是 G 的一个子群, 则 $s\mathrm{A}s^{-1}$ 也是 G 的一个子群 (称它是 A 在 G 内的一个**共轭子群**).

称使得 $s\mathrm{A}s^{-1} = \mathrm{A}$ 的 $s \in \mathrm{G}$ 的集合 N(A) 为 G 的子群 A 的**正规化子**, 它显然也是 G 的子群. 说明 A 的中心化子 (习题 11) 是 A 的正规化子的一个不变子群.

14. 设 G 是作用在集合 X 上的群.

a) 说明关系

$$\text{存在一个 } s \in \mathrm{G}, \quad \text{使得 } y = sx$$

是集合 X 上的一个等价关系 (一个元素 $x \in \mathrm{X}$ 关于这个关系的等价类称为被 G 作用 x 的**轨道**). 说明对于所有 $x \in \mathrm{X}$, 使得 $sx = x$ 的 $s \in \mathrm{G}$ 的集合是 G 的一个子群 H_x (称为 x 在 G 内的**稳定化子**), 并且同一个轨道的不同的元素的稳定化子在习题 13 的意义下是在 G 内相互共轭的.

b) 对于一个元素 $x \in \mathrm{X}$ 考虑由

$$f(s) = sx$$

给定的从 G 到 X 内的映射 f. 说明 f 是从 G 到 G/H_x 上的典范映射和从 G/H_x 到 X 内的一个映射的复合, 并说明后一个映射诱导一个从 G/H_x 到 x 被 G 作用的轨道上的双射, 并且如果 G 是有限的, 则

$$\mathrm{Card(G)} = \mathrm{Card(M)} \cdot \mathrm{Card}(\mathrm{H}_x).$$

c) 如果 X 为通常的空间, 而 G 是绕 X 内给定点 O 旋转的群, 描述轨道和稳定化子.

¶d) 假定 G 是阶为素数 p 的一个幂的有限群, 而 X 是有限的, 其元素的数目不是 p 的倍数. 说明 G 在 X 内至少具有一个**不动点** (即存在一个 $x \in \mathrm{X}$, 使得对于所有 $s \in \mathrm{G}$ 有 $sx = x$).

¶e) 假定 G 是一个 p **群**, 即其元素个数是一个素数 p 的幂. 令 G 通过内自同构作用在自身上 (例 19), 说明 G 的中心 (习题 11) 不会缩减为一个中性元.

¶15. 设 G 是一个群, 而 H 是 G 的一个子群. 让 G 作用在 G/H 上 (例 20). 说明 G/H 的其稳定化子包含 H 的元素是子群 N(H) 的元素在从 G 到 G/H 上的典范映射下的像, 这里 N(H) 是上面的习题 13 中定义的 H 在 G 内的正规化子.

¶16. 设 H 是 G 的一个不变子群. 说明在 G/H 上存在唯一的一个运算, 它使得 G/H 成为一个群, 并且从 G 到 G/H 内的典范映射是群的一个同态 (利用 §4 的定理 3); 这样得到的群称为 G 关于 H 的**商群**. 如果取有理整数的加法群 **Z** 作为 G, 而 H 是 G 的一个子群, 将会发生什么?

设 p 是从 G 到 G/H 的典范映射, 说明对于 G/H 的所有子群 A, 存在唯一的一个 G 的包含 H 的子群 K, 使得 $\mathrm{A} = p(\mathrm{K})$, 并且还有 $\mathrm{K} = p^{-1}(\mathrm{A})$.

称由 G 的形式为 $xyx^{-1}y^{-1}$ 的元素生成的子群为 G 的**导子群**, 记为 G′ 或 D(G). 说明 D(G) 是 G 的不变子群, 并且如果 H 是 G 的一个不变子群, 那么商群 G/H 是交换的, 必须并且只需 $\mathrm{H}\supset \mathrm{D(G)}$.

¶¶17. 给定群 G 的两个子群 A 和 B, 记由 $xyx^{-1}y^{-1}$, 其中 $x\in \mathrm{A}$ 且 $y\in \mathrm{B}$, 生成的 G 的子群为 (A, B). 令

$$\mathrm{D(G)=(G,G)},\quad \mathrm{D^2(G)=D(D(G))},\quad \mathrm{D^3(G)=D(D)^2(G))},\quad \cdots,$$

说明下列条件是等价的:

a) 存在一个整数 r, 使得 $\mathrm{D}^{r+1}(\mathrm{G})=\{e\}$.

b) 可以构造 G 的子群

$$\{e\}=\mathrm{H}_0\subset \mathrm{H}_1\subset\cdots\subset \mathrm{H}_s=\mathrm{G},$$

使得对于满足条件 $0\leqslant i\leqslant s-1$ 的每个指标 i, 子群 H_i 在 H_{i+1} 内是不变的, 并且商群 (商群在习题 16 中定义)$\mathrm{H}_{i+1}/\mathrm{H}_i$ 是交换的.

c) 可以构造 G 的不变子群

$$\{e\}=\mathrm{K}_0\subset \mathrm{K}_1\subset\cdots\subset \mathrm{K}_s=\mathrm{G},$$

使得所有商群 $\mathrm{K}_{i+1}/\mathrm{K}_i$ 是交换的.

满足这些条件的群称为**可解的**. 证明一个可解群的所有子群是可解的.

设 G 是一个群, 而 H 是 G 的一个不变子群. 如果群 H 和 G/H 是可解的, 则 G 也是可解的.

18. 设 G 是一个 p 群 (习题 14). 证明 G 的所有子群和所有商群是一个 p 群. 利用习题 14 的 e) 证明 G 包含不变子群

$$\{e\}=\mathrm{H}_0\subset \mathrm{H}_1\subset\cdots\subset \mathrm{H}_s=\mathrm{G},$$

使得对于 $1\leqslant i\leqslant r$, 商群 $\mathrm{H}_i/\mathrm{H}_{i-1}$ 同构于加法群 $\mathbf{Z}/p\mathbf{Z}$ (即是 p 阶循环群). 一个特殊推论是 p 群为可解的.

19. 设 G 是 n 阶有限循环群.

a) 证明对于 n 的所有因数 d, 使得 $x^d=e$ 的 $x\in \mathrm{G}$ 的个数是 d.

b) 设 d 是 n 的一个因数. 一个 $x\in \mathrm{G}$ 可以对于一个适当选取的 $y\in \mathrm{G}$ 写成形式 y^d, 必须并且只需

$$x^{n/d}=e.$$

¶¶20. 设 G 是一个 n 阶有限交换群. 假定对于 n 的所有因数 d, 使得

$$x^d=e$$

的 $x\in \mathrm{G}$ 至多有 d 个. 我们打算由此推出 G 是循环的 (对于有限域的研究, 这个结果是不可缺少的: 参见 §33, 习题 2). 在下面, 用

$$n=p_1^{r_1}\cdots p_h^{r_h}$$

表示 n 的素因数分解.

a) 证明对于所有满足条件 $1 \leqslant i \leqslant h$ 的 i, 存在一个 $a_i \in \mathrm{G}$, 满足条件

$$a_i^{p_i^{r_i}} = e, \quad a_i^{p_i^{r_i-1}} \neq e,$$

并且 a_i 的阶恰好是

$$q_i = p_i^{r_i}.$$

b) 利用 q_i 两两互素的事实说明元素 $a_1 \cdots a_h$ 是 $q_1 \cdots q_h = n$ 阶的, 并且 G 如所宣布的那样是循环的.

c) 把假设减弱为: 对于 $1 \leqslant i \leqslant h$, 使得

$$x^{p_i} = e$$

的 $x \in \mathrm{G}$ 的数目最多是 p_i, 说明可以得到同一个结论.

21. 设 f 是从一个有限群 G 到一个群 H 的同态. 说明

$$\mathrm{Card}(\mathrm{G}) = \mathrm{Card}(\mathrm{Ker}(f))\mathrm{Card}(\mathrm{Im}(f)).$$

22. 设 G 是一个有限交换群, 而 n 是这样的整数(*)

$$x^n = e \quad 对于所有的\ x \in \mathrm{G}.$$

a) 假定 $n = rs$, 这里 r 和 s 互素. 设 M (对应的, N) 是使得

$$x^r = e\ (对应的, x^s = e)$$

的 $x \in \mathrm{G}$ 的集合. 说明 M 和 N 是 G 的子群. 通过写出对于 r 和 s 的 Bezout 等式, 证明由 $f(x, y) = xy$ 给定的映射

$$f : \mathrm{M} \times \mathrm{M} \to \mathrm{G}$$

是群的一个同构.

b) 设

$$n = p_1^{r_1} \cdots p_h^{r_h} = q_1 \cdots q_h, \quad 其中\ q_i = p_i^{r_i}$$

是 n 的素因数分解; 对于满足条件 $1 \leqslant i \leqslant h$ 的所有的 i, 设 M_i 是使得

$$x^{q_i} = e$$

的 $x \in \mathrm{G}$ 的元素组成的子群. 证明 G 同构于群 $\mathrm{M}_1, \cdots, \mathrm{M}_h$ 的直积.

c) 设 M 是一个有限交换群, p 是一个素数, 而 r 是一个非负整数; 假定对于所有 $x \in \mathrm{M}$ 有

$$x^{p^r} = e,$$

证明 Card(M) 是 p 的一个幂 (观察到如果 $\mathrm{M} \neq \{e\}$, 则可以找到 M 的一个 p 阶子群 M′; 引进商群 M/M′, 就可以关于 M 的元素个数进行归纳推理).

(*) 这个习题的陈述是按照乘法记号书写的, 但是读者把它翻译成加法记号的陈述对于今后的推广是有益的.

d) 证明下列定理: 设 G 是阶为

$$n = p_1^{r_1} \cdots p_h^{r_h}$$

的有限交换群, 则 G 同构于阶分别为 $p_1^{r_1}, \cdots, p_h^{r_h}$ 的 h 个群的直积 (这个结果在 1801 年 Gauss 已经给了差不多的证明, 在 §31 的习题中将会得到补充: 其阶为 p 的幂的交换群同构于其阶为 p 的幂的*循环群*的直积. *有限生成的交换群*的完整研究由 Kronecker 在 1870 年做出; 一个这样的群是一个有限交换群和一个群 $\mathbf{Z}^n$ 的直积).

¶23. 再回到习题 22 的 a). 设 A (对应的, B) 是由对于某个 $y \in \mathrm{G}$ 使得

$$x = y^s \text{ (对应的}, x = y^r)$$

的 $x \in \mathrm{G}$ 组成的 G 的子群. 说明 A = M 和 B = N (证明有关系

$$\mathrm{Card(G)} = \mathrm{Card(M)} \cdot \mathrm{Card(N)} = \mathrm{Card(A)} \cdot \mathrm{Card(N)} = \mathrm{Card(B)} \cdot \mathrm{Card(M)},$$

并且观察到 $\mathrm{A} \subset \mathrm{M}, \mathrm{B} \subset \mathrm{N}$).

假定 $n = \mathrm{Card(G)}$. 证明 $\mathrm{Card(M)} = r, \mathrm{Card(N)} = s$.

24. 设 $s \in \mathfrak{S}_n$ 是 $\mathrm{X} = \{1, 2, \cdots, n\}$ 的一个置换, 而 G 是由 s 的幂组成的 $\mathfrak{S}_n$ 的子群.

a) 说明可以找到 X 的满足下列条件的非空子集 $\mathrm{I}_1, \cdots, \mathrm{I}_r$: 对于 $1 \leqslant k \leqslant r$ 和所有 $g \in \mathrm{G}$ 有 $g(\mathrm{I}_k) = \mathrm{I}_k$; 集合 I_k 是两两不交的, 并且其并集为整个 X; $p, q \in \mathrm{X}$ 属于同一个 I_k, 必须并且只需存在一个 $g \in \mathrm{G}$, 使得 $q = g(p)$. 与习题 14 的 a) 的关系如何?

b) 证明前面的条件刻画了集合 I_k 的特征 (自然可以调整写出它们的次序).

c) 证明对于每个 k, 可以写出 I_k 的元素, 其形式是一个序列 $i_0, \cdots, i_p$, 使得

$$s(i_0) = i_1, s(i_1) = i_2, \cdots, s(i_{p-1}) = i_p, s(i_p) = i_0$$

(*一个置换的轮换分解*; 称写成自然顺序的. 两两不同的整数的所有满足上述关系的序列 $i_0, \cdots, i_p$ 为 s 的一个**轮换**).

d) 找出下列置换的轮换:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 1 & 3 & 6 & 5 & 7 & 4 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 8 & 9 & 2 & 1 & 4 & 3 & 6 & 7 \end{pmatrix}.$$

(注意, 在这里我们使用表示一个置换 s 的标准记号. 这个记号的组成是在第二行写出第一行对应元素的像.)

e) 给定一个置换 $s \in \mathfrak{S}_n$, 设 $n_1, \cdots, n_r$ 是 s 的不同循环的数目. 证明 s 在 $\mathfrak{S}_n$ 中的阶 (即使得 $s^q = e$ 的最小整数 $q \geqslant 1$, 或由 s 生成的循环群的阶) 是整数 $n_1, \cdots, n_r$ 的最小公倍数.

f) 考虑置换

$$s : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 5 & 4 & 1 & 7 & 10 & 2 & 6 & 9 & 8 \end{pmatrix},$$

计算 $\mathfrak{S}_{10}$ 的由 s 生成的子群的阶. 计算置换 s^{100}.

¶¶25. 在这个习题里, 利用以下术语. 给定由加法群 G_i 和同态 $f_i : G_n \to G_{n+1}$ 组成的一个序列

$$G_1 \xrightarrow{f_1} G_2 \xrightarrow{f_2} G_3 \longrightarrow \cdots \longrightarrow G_n \xrightarrow{f_n} G_{n+1},$$

说这个序列是**恰当的**, 如果对于每个满足条件 $1 \leqslant i < n$ 的 i, 同态 $f_i : G_i \to G_{i+1}$ 的像是下一个同态 f_{i+1} 的核. 此外称一个由集合和从这些集合到另一些集合的映射组成的图, 例如

$$\begin{array}{ccccccccc} A & \xrightarrow{f} & B & \xrightarrow{g} & C & \xrightarrow{h} & D & \xrightarrow{k} & E \\ \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle q} & & \downarrow{\scriptstyle r} & & \downarrow{\scriptstyle s} & & \downarrow{\scriptstyle t} \\ A' & \xrightarrow{f'} & B' & \xrightarrow{g'} & C' & \xrightarrow{h'} & D' & \xrightarrow{k'} & E' \end{array} \tag{1}$$

是**交换的**, 如果对于图的任何 “顶点”X 和 Y, 所有在图中复合得到的从 X 到 Y 的映射是相等的. 在图 (1) 的情形, 交换性翻译成关系 $f' \circ p = q \circ f$ 和许多其他的类似关系, 读者将会写出.

我们考虑一个交换图 (1), 其中的 $A, B, \cdots, E'$ 是加法群, 映射是群的同态映射. 假定图的水平线是恰当的, 由此得到

$$\mathrm{Im}(f) = \mathrm{Ker}(g), \quad \mathrm{Im}(f') = \mathrm{Ker}(g'),$$

等等, 建立下列结果 (以**五引理**而著名):

a) 如果 p 是满射的, 而 q 和 s 是单射的, 则 r 是单射的.

b) 如果 q 和 s 是满射的, 而 t 是单射的, 则 r 是满射的.

c) 如果 p 是满射的, q 和 s 是满射的, 而 t 是单射的, 则 r 是双射的.

§8 环和域

1. 环的定义, 例子

称由一个集合 K 和 K 上的记作 $(x, y) \to x + y$ (K 内的 “加法”) 和 $(x, y) \to xy$ (K 内的 “乘法”) 的两个运算组成的三元组为一个**环**, 如果它们满足下列条件:

(A1) 集合 K 和 K 上的运算 $(x, y) \to x + y$ 组成的序偶是一个交换群;

(A2) 运算 $(x, y) \to xy$ 是结合的, 并且有一个中性元(*);

(A3) 对于任意 $x, y, z \in K$ 有关系 (乘法对于加法的分配律)

$$x(y + z) = xy + xz, \quad (x + y)z = xz + yz.$$

环的公理 (A1) 断言如下: 首先有等式

$$x + (y + z) = (x + y) + z,$$

$$x + y = y + x;$$

(*) 一些作者在环的定义中不要求乘法的中性元的存在性, 这样就得到一个比本书更一般的环的概念. 但是实际情形表明, 并且理论上也证实, 像本著作所做的那样, 总可以限于研究带单位元的环.

其次, 在 K 内存在一个元素, 记为 0, 使得对于所有 $x \in \mathrm{K}$ 有

$$x + 0 = x;$$

最后, 对于任意 $x \in \mathrm{K}$, 存在 K 的一个元素, 记为 $-x$, 使得

$$x + (-x) = 0.$$

由上推出, 对于任意 $a, b \in \mathrm{K}$, 方程

$$a + x = b$$

有且仅有一个解, 即 $b + (-a)$, 记为

$$x = b - a.$$

环的公理 (A2) 意味着在一个环里, 对所有 $x, y, z \in \mathrm{K}$, 有等式

$$x(yz) = (xy)z,$$

并且存在 K 的一个元素, 记为 1, 使得对于所有 $x \in \mathrm{K}$ 有

$$x1 = 1x = x.$$

我们称 1 是 K 的**单位元**.

我们说环 K 是**交换的**, 如果

$$xy = yx \quad \text{对于任意的 } x, y \in \mathrm{K}.$$

在矩阵论里将会遇到非交换的环的例子. 我们说环 K 中的两个元素 x, y 是**交换的**, 如果有 $xy = yx$.

在一个环里, 对于所有 x 有关系

$$(-1)x = -x, \quad 0x = 0.$$

为了证明第一个关系, 只需证明 $(-1)x + x = 0$, 而

$$(-1)x + x = (-1)x + 1x = (-1 + 1)x = 0x,$$

所以归结为证明第二个关系 $0x = 0$. 为此我们计算

$$0x + x = 0x + 1x = (0 + 1)x = 1x = x,$$

这表明 $0x = x - x = 0$, 这在所有加法群里是成立的.

现在举几个环的例子.

例 1 给有理整数集 $\mathbf{Z}$ 配备通常的两个运算 (加法和乘法), 显然得到一个交换环, 称为**有理整数环**. 配备了通常的加法和乘法的集合 $\mathbf{Q}$ (有理数集合) 和 $\mathbf{R}$ (实数集合) 也是交换环.

例 2 设 K 是具有下列性质的实数集合: 存在有理整数 a, b, c, 使得

$$x = a + br + cr^2, \text{ 其中 } r = \sqrt[3]{2}.$$

可以直接验证, 如果 $x, y \in \mathrm{K}$, 那么 $x + y$ 和 xy 仍在 K 内. 这就允许给 K 配备通常的加法和乘法运算. 因此, K 和这两个运算一起构成一个交换环.

例 3 设 X 是任意一个集合, K 是任意一个环, 用 A 表示从 X 到 K 内的所有映射的集合. 对于 $f, g \in \mathrm{A}$, 我们如下定义 $f + g$ 和 fg: $f + g$ 是从 X 到 K 内的映射

$$x \to f(x) + g(x),$$

而 fg 是从 X 到 K 内的映射

$$x \to f(x)g(x).$$

由此, 集合 A 配备了刚定义的这两个运算

$$(f, g) \to f + g \quad \text{和} \quad (fg) \to fg$$

就是一个环 (当且仅当 K 是交换的, 它是交换的). 例如验证分配律

$$f(g + h) = fg + fh;$$

只需验证等式两端 (它们是从 X 到 K 内的映射) 在每个 $x \in \mathrm{X}$ 有相同的值: 左端的值是 $f(x)(g(x) + h(x))$, 右端的值是 $f(x)g(x) + f(x)h(x)$. 我们看到在 A 内分配律是满足的, 因为在 K 内已经满足.

建议初学者详细处理这个例子, 并且确信这样一个事实, 为了指出 A 是一个环, 必需援引 K 满足环的所有公理. 考察 X 是仅有一个元素的集合就可以容易地明白为什么必须这样.

刚定义的环称为**从集合 X 到环 K 内的映射环**.

例 4 在例 3 中取 $\mathrm{X} = \mathrm{K} = \mathbf{R}$, 但不是考虑所有从 X 到 K 内的映射, 而只考虑满足某些事先给定的 "正则性" 条件的映射, 比如, 在一个给定点是连续的, 是处处连续的, 处处有三阶连续导数, 等等; 在每一种情形都得到一个交换环.

例 5 在后面 §15 我们将看到系数在一个给定的环 K 内的 n 行 n 列的方阵组成一个新的环, 只要按照 §14 和 §15 中的公式定义矩阵的加法和乘法.

设 K 是一个环, 称 K 的所有满足下列条件的子集 A 是 K 的**子环**: A 是加法群 K 的一个子群; 关系 $x \in A$ 和 $y \in A$ 蕴含关系 $xy \in A$; 并且 $1 \in A$. 如果是这样, 那么从 $K \times K$ 到 K 的映射

$$(x, y) \to x + y \quad 和 \quad (x, y) \to xy$$

映射 $A \times A$ 到 A 内, 从而定义了集合 A 上的两个运算. 由此, 配备了这两个运算的集合 A 是一个环.

事实上, 首先, 根据 §7 第 3 小节配备了加法的集合 A 是一个交换群; 其次, 乘法在 K 内是结合的, 在 A 内更是如此, 并且 A 有一个关于乘法的中性元, 因为 A 包含 K 的中性元; 最后, 分配律在 K 内成立, 在 A 内更有理由成立.

显然在上面的例 1 中, $\mathbf{Z}$ 是 $\mathbf{Q}$ 的一个子环, $\mathbf{Q}$ 自身是 $\mathbf{R}$ 的一个子环. 此外, 上面例 4 中的环是所有从集合 $X = \mathbf{R}$ 到环 $K = \mathbf{R}$ 内的映射的环的子环.

我们注意到为了验证环 K 的一个子集 A 是 K 的子环, 只需验证下列条件: 如果 A 含有两个元素 x 和 y, 则它也含有它们的和 $x + y$ 和它们的乘积 xy, 并且 A 含有 -1.

事实上, 假定条件满足. 对于 $x \in A$, 考虑到 $-x = (-1)x$, 我们有 $-x \in A$. 如果 A 含有 x 和 y, 则也含有 x 和 $-y$, 于是含有 $x + (-y) = x - y$, 这就证明了 A 是加法群 K 的一个子群, 换句话说, 出现在子环定义中的第一个条件满足. 再有, 由于 A 含有 -1, 它应该含有 $-(-1) = 1$, 从而子环定义中的第三个条件满足. 最后, 根据假设第二个条件满足.

2. 整环和域

在一个环 K 内考虑方程

$$ax = b, \tag{1}$$

其中 a, b 是给定的元素, 而 x 是 K 的一个 "未知" 元素.

第一个简单的情形是 $a = 0$. 由于对于所有 $x \in K$ 有 $0x = 0$, 那么显然仅有两种可能的情形: 或者 $b = 0$, 这时所有的 $x \in K$ 都满足 (1); 或者 $b \neq 0$, 那么方程 (1) 没有任何解.

第二个十分简单的情形是 a 具有关于乘法的一个逆元, 即存在 K 的一个 (且仅一个) 元素, 记作 a^{-1}, 满足

$$a^{-1}a = aa^{-1} = 1;$$

那么 §6 的定理 4 适用: 对于任意 $b \in K$, 方程 (1) 有且仅有一个解

$$x = a^{-1}b.$$

在这种情形, 我们称 a 是 K 的一个**可逆元**或**单位**.

在所有可能的情形中, 还要考察余下的一种可能情形, a 既不是 0, 也不是可逆的. 首先, 最好不过的是这种情形不发生. 换句话说, 可能出现 K 的所有非零元素都是可逆的这种情形, 这时就说 K 是一个域(*). $\mathbf{Q}$ 和 $\mathbf{R}$ 是域 (**有理数域**和**实数域**). 反之, 环 $\mathbf{Z}$ 不是一个域 ($x \in \mathbf{Z}$ 在环 $\mathbf{Z}$ 内是可逆的, 必须且只需存在一个 $y \in \mathbf{Z}$ 使得 $xy = 1$, 显然只有 $x = 1$ 或 $x = -1$ 时这才是可能的).

回到一般情形, 可以问方程 (1) 是否可能具有多个解. 如果 x 和 y 是两个这样的解, 显然有 $ax = ay$, 于是

$$a(x-y) = 0,$$

这就导致引进下列概念: 称一个环 K 是**整**的或称 K 是**整环**, 如果对于 $u \in \mathrm{K}$ 和 $v \in \mathrm{K}$,

$$\text{关系 } uv = 0 \text{ 蕴含 } u = 0 \text{ 或 } v = 0.$$

(换句话说, 如果两个因子都非零, 则其乘积非零.) 环 $\mathbf{Z}$ 显然是一个整环. 域必然是整环, 因为如果 $u \neq 0$, 关系 $uv = 0$ 推出 $v = u^{-1}0 = 0$.

在一个整环内, 方程 (1) 至多有一个解, 正如上面的推理所表明的那样. 但是自然可能发生它根本就没有解 —— 比如在环 $\mathbf{Z}$ 里方程 $2x = 3$ 就是这种情形 —— 一般, 在不是域的整环的情形, 关于方程 (1) 可解的条件不能得到任何结论.

存在非整的环. 作为例子, 取从 $\mathbf{R}$ 到环 $\mathbf{R}$ 内的所有映射的环 (上面的例 3), 并且考虑这个环的两个如下定义的元素:

$$f(x) = \begin{cases} x, & x \geqslant 0, \\ 0, & x \leqslant 0, \end{cases} \qquad g(x) = \begin{cases} 0, & x \geqslant 0, \\ x, & x \leqslant 0; \end{cases}$$

显然有

$$f(x)g(x) = 0 \quad \text{对于所有 } x \in \mathbf{R}.$$

从而在所考虑的环里 $fg = 0$; 但是 $f \neq 0$, 并且 $g \neq 0$ (因为从一个集合 X 到一个环 K 内的映射的环的零元是函数, 在每个 $x \in \mathrm{X}$ 无例外地取 0 值, 而不论 f 也不论 g 都不是这种情形).

注 1 设 K 是一个环, 习惯上用记号



$$\mathrm{K}^*$$

表示 K 的可逆元素的集合 (参见已使用过的 $\mathbf{Q}^*$ 和 $\mathbf{R}^*$ 的符号). 根据 §6 的定理 3, 如果 K^* 含有两个元素 x 和 y, 则它也含有 xy. 于是可以给 K^* 配备运算 $(x, y) \to xy$. 如此说来, 集合 K^* 配备了这个运算后是一个群. 事实上, 首先显然 K^* 上的运算是结合的 (因为在 K 内已经是这样的); 其次显然有 $1 \in \mathrm{K}^*$, 所

(*) 事实上, 在一个域内, 还要求 $1 \neq 0$, 以致一个域至少总具有两个元素.
本书中的 "域" 是我们代数书中的 "除环"; 本书中的 "交换域" 是我们的 "域". —— 译者注

以集合 K^* 上的运算具有中性元; 最后, 如果 $x \in K^*$, 根据 §6 的定理 3, 也有 $x^{-1} \in K^*$, 而由于我们有

$$x^{-1}x = xx^{-1} = 1,$$

1 是 K^* 的中性元, 我们看到 K^* 的所有元素对于所考虑的运算是可逆的 (在 K^* 内, 不仅在 K 内).

配备了运算 $(x, y) \to xy$ 的 K^* 称为**环 K 的乘法群**. 如果 K 是一个域, 我们有

$$K^* = K - \{0\}.$$

作为一个特殊例子, 我们有

$$\mathbf{Z}^* = \{1, -1\},$$

其乘法表是

$$1 \cdot 1 = 1; \quad (-1) \cdot 1 = 1 \cdot (-1) = -1; \quad (-1) \cdot (-1) = 1.$$

注 2 设 K 是一个域. 称 K 的所有满足下列条件的子集 A 为 K 的**子域**: A 是 K 的一个子环, 并且对于 $x \neq 0$, 关系 $x \in A$ 蕴含 $x^{-1} \in A$. 那么显然配备了由 K 的运算所 "诱导的" 运算的集合 A 不仅是一个环, 而且是一个域.

这样看来, $\mathbf{Q}$ 是 $\mathbf{R}$ 的一个子域.

例 6 设 $K \subset \mathbf{R}$ 是具有下列性质的实数 x 的集合: 存在有理数 a 和 b 使得

$$x = a + br, \text{ 其中 } r = \sqrt{2}.$$

读者容易验证 K 是 $\mathbf{R}$ 的一个子域.

3. 模 p 整数环

在 §4 中的例 9, 我们对于所有有理整数 p 定义了模 p 整数的集合 $\mathbf{Z}/p\mathbf{Z}$, 并且在同一节的例 14 中我们在这个集合上定义了称为加法和乘法的两个运算, 这两个运算跟在普通整数上的运算按下列关系相联系: 如果 θ 表示从 $\mathbf{Z}$ 到 $\mathbf{Z}/p\mathbf{Z}$ 上的典范映射, 则对于任意 $x, y \in \mathbf{Z}$, 有

$$\theta(x + y) = \theta(x) + \theta(y), \quad \theta(xy) = \theta(x)\theta(y).$$

由此, 我们将推出配备了 §4 定义的加法和乘法的模 p 整数的集合 $\mathbf{Z}/p\mathbf{Z}$ 是一个交换环.

首先指出在 $\mathbf{Z}/p\mathbf{Z}$ 里加法是结合的: 设 ξ, η, ς 是这个集合的三个元素, 存在 $x, y, z \in \mathbf{Z}$ 使得 $\xi = \theta(x), \eta = \theta(y), \varsigma = \theta(z)$, 我们有

$$\xi + \eta = \theta(x) + \theta(y) = \theta(x + y), \quad \eta + \varsigma = \theta(y) + \theta(z) = \theta(y + z),$$

于是

$$(\xi+\eta)+\varsigma=\theta(x+y)+\theta(z)=\theta((x+y)+z),$$
$$\xi+(\eta+\varsigma)=\theta(x)+\theta(y+z)=\theta(x+(y+z)),$$

从而由 $\mathbf{Z}$ 的加法的结合性推出 $\mathbf{Z}/p\mathbf{Z}$ 的加法的结合性.

同样证明在 $\mathbf{Z}/p\mathbf{Z}$ 里乘法的结合性, 加法和乘法的交换性以及乘法对于加法的分配性.

从关系

$$\theta(1)\theta(x)=\theta(1x)=\theta(x)$$

看出 $\theta(1)$ 显然是在 $\mathbf{Z}/p\mathbf{Z}$ 里乘法的中性元, 同样 $\theta(0)$ 是加法的中性元.

为了证明 $\mathbf{Z}/p\mathbf{Z}$ 是一个交换环, 尚需指出 $\mathbf{Z}/p\mathbf{Z}$ 的所有元素 ξ 有一个相反元; 为此我们适当选取一个 $x\in\mathbf{Z}$ 写出 $\xi=\theta(x)$, 考虑到关系

$$\theta(-x)+\theta(x)=\theta(-x+x)=\theta(0),$$

显然得到 ξ 有相反元, 这就是 $\theta(-x)$.

我们刚证明的结果允许我们谈论**模 p 整数环** $\mathbf{Z}/p\mathbf{Z}$. 如果 $p\neq 0$, 这个环仅有有限个元素, 即 p 个元素 (可以假定 p 是正的, 这不失一般性).

对于 p 的某些值, 环 $\mathbf{Z}/p\mathbf{Z}$ 甚至是一个域 (这将证明**有限域**的存在性, 所谓有限域是具有有限个元素的域):

定理 1 设整数 $p\geqslant 2$, 以下断言等价:

a) $\mathbf{Z}/p\mathbf{Z}$ 是整环;

b) $\mathbf{Z}/p\mathbf{Z}$ 是一个域;

c) p 是素数.

设 ξ 和 η 是 $\mathbf{Z}/p\mathbf{Z}$ 的两个非零元; 我们有 $\xi=\theta(x)$, $\eta=\theta(y)$, 其中

$$x\not\equiv 0(\text{mod } p),\quad y\not\equiv 0(\text{mod } p);$$

为了由此推出 $\xi\eta=\theta(xy)$ 也是非零元, 必须证明

$$xy\not\equiv 0(\text{mod } p).$$

换句话说, $\mathbf{Z}/p\mathbf{Z}$ 是整环, 必须且只需

$$\text{关系 } xy\equiv 0(\text{mod } p) \text{ 蕴含 } x\equiv 0(\text{mod } p) \text{ 或 } y\equiv 0(\text{mod } p),$$

或者等价地说, 如果 p 整除一个乘积 xy, 则它整除 x 或整除 y, 由此得到性质 a) 和 c) 等价.

b) 显然蕴含 a). 为了完成证明, 只需证明 a) 蕴含 b), 这明显地由更一般的下列定理推出:

定理 2 所有有限整环是一个域.

设 K 是一个有限整环, 对于 K 的一个元素 $a \neq 0$, 考虑从 K 到 K 内的映射 $x \to ax$. 由于 $ax = ay$ 蕴含 $a(x-y)=0$, 故如果 K 是整环, 必有 $x - y = 0$, 我们看到所考虑的映射是单射的. 但由于 K 是有限的, 故这个映射是满射的 (§5, 定理 4), 特殊情形是 $ax = 1$ 可解, 这就表明所有非零元素有右逆元. 借助映射 $x \to xa$ 同样证明 K 的所有非零元有左逆元, 这就完成了证明.



注 3 可以证明所有有限域是交换的, 但是为了证明这个结论所必需的技术远远超过了本书的水平.

此外, 还可以证明有限域的元素个数必定是一个素数的幂, 并且对于所有素数 p 和所有整数 $n \geqslant 1$, 本质上仅存在一个有 p^n 个元素的域 (这说明我们知道怎样明晰地构造所有的有限域). 有限域的最初的详细研究归功于 Galois.

4. 二项式公式

在中学证明的大部分 "代数恒等式" 在所有的环里仍然有效 (有时要假定出现在等式里的元素 $x, y, \cdots$ 是两两交换的). 比如, 设 x, y 是一个环 K 里的两个元素, 我们做计算

$$(x+y)^2 = (x+y)(x+y) = x^2 + xy + yx + y^2;$$

我们看到, 如果 x 和 y 是交换的, 就有等式 $(x+y)^2 = x^2 + 2xy + y^2$. 于是经过平凡的计算即得

$$(x+y)^3 = (x^2+2xy+y^2)(x+y) = x^3 + 3x^2y + 3xy^2 + y^3.$$

更一般的, 有

定理 3 设 K 是一个环, x 和 y 是 K 的两个元素, 并且假定 x 和 y 是交换的. 则对于所有整数 $n \geqslant 1$ 有

$$(x+y)^n = \sum_{p=0}^{n} \binom{n}{p} x^{n-p} y^p,$$

其中

$$\binom{n}{p} = \frac{n(n-1)\cdots(n-p+1)}{1 \cdot 2 \cdots p} = \frac{n!}{p!(n-p)!} \quad (0 \leqslant p \leqslant n).$$



注 4 我们提醒 (§5, 定理 10) 数 $\binom{n}{p}$ 是正整数, 而不仅仅是有理数.

所要证明的结果对于 $n = 1$ 是平凡的, 只需证明等式

$$(x+y)^{n-1} = \sum_{p=0}^{n-1} \binom{n-1}{p} x^{n-1-p} y^p$$

蕴含对于指数 n 的类似公式. 而当上式两端乘以 $x+y$, 便得

$$(x+y)^n = \sum_{p=0}^{n-1}\binom{n-1}{p}x^{n-p}y^p + \sum_{p=0}^{n-1}\binom{n-1}{p}x^{n-1-p}y^{p+1};$$

如果整数 r 满足条件 $0<r<n$, 那么上式右端有两项含有单项式

$$x^{n-r}y^r,$$

第一项由在第一个和式中取 $p=r$ 得到, 这引进了一个因子, 它等于

$$\binom{n-1}{r},$$

而第二项由在第二个和式中取 $p=r-1$ 得到, 这引进了一个因子, 它等于

$$\binom{n-1}{r-1}.$$

为了完成证明还剩下验证关系

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}.$$

其右端等于

$$\begin{aligned}
&\frac{(n-1)(n-2)\cdots(n-r)}{1\cdot 2\cdots r} + \frac{(n-1)(n-2)\cdots(n-r+1)}{1\cdot 2\cdots(r-1)}\\
=&\frac{(n-1)(n-2)\cdots(n-r)+(n-1)(n-2)\cdots(n-r+1)r}{1\cdot 2\cdots r}\\
=&\frac{[(n-r)+r](n-1)(n-2)\cdots(n-r+1)}{1\cdot 2\cdots r}\\
=&\frac{n(n-1)(n-2)\cdots(n-r+1)}{1\cdot 2\cdots r} = \binom{n}{r}
\end{aligned}$$

这就完成了证明.

关系

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$$

使得可以容易地计算 $\binom{n}{r}$, 基于明显的理由, 称为**二项式系数**. 这些系数由下表给

定, 此表称为 **Pascal 三角形**:

$$
\begin{array}{cccccccc}
1 & 1 & & & & & & \\
1 & 2 & 1 & & & & & \\
1 & 3 & 3 & 1 & & & & \\
1 & 4 & 6 & 4 & 1 & & & \\
1 & 5 & 10 & 10 & 5 & 1 & & \\
1 & 6 & 15 & 20 & 15 & 6 & 1 & \\
1 & 7 & 21 & 35 & 36 & 21 & 7 & 1 \\
\end{array}
$$

$$\cdots\cdots\cdots\cdots$$

计算第 n 行的第 p 项的方法是, 把前一行的第 $p-1$ 项和第 p 项相加.

观察上表启发我们有等式

$$\binom{n}{r} = \binom{n}{n-r};$$

由于

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}, \quad \binom{n}{n-r} = \frac{n!}{(n-r)!r!};$$

上述等式的验证是容易的. 但是所论等式的真正的理由在于事实: 考虑到表达式

$$(x+y)^n = \sum_{p=0}^{n} \binom{n}{p} x^{n-p} y^p$$

的意义, 置换其中的 x 和 y 应当是不变的; 于是这个表达式中 "单项式" $x^{n-y}y^r$ 和 $x^r y^{n-r}$ 的系数自然应当相等, 因为这两个单项式在 x 和 y 互换时一个变成另一个.

还可以做如下观察, 用 X 表示 n 个元素的集合, $\binom{n}{r}$ 是 X 的含有 r 个元素的子集的个数, 令每个这样的子集 Y 对应它的补集 X − Y, 就得到从 X 的含有 r 个元素的子集的集合到 X 的含有 $n-r$ 个元素的子集的集合上的一个双射, 由此即得

$$\binom{n}{r} = \binom{n}{n-r}.$$

5. 和的乘积展开

二项式公式是一个更一般的公式的特例, 我们来介绍这个一般公式.

设 K 是一个交换环, I 是一个有限集合, 而 $(x_i)_{i\in \mathrm{I}}$ 和 $(y_i)_{i\in \mathrm{I}}$ 是以 I 为指标的 K 的元素的两个族. 我们要 "展开" 乘积

$$\prod_{i\in \mathrm{I}} (x_i + y_i).$$

为了陈述结果, 首先引进下列记号: 给定 I 的一个子集 F, 我们令

$$x_{\mathrm{F}} = \prod_{i\in\mathrm{F}} x_i, \quad y_{\mathrm{F}} = \prod_{i\in\mathrm{F}} y_i$$

(当然如果 F 是空集, 则 $x_{\mathrm{F}} = y_{\mathrm{F}} = 1$). 然后, 要找的公式就可以写成

$$\prod_{i\in\mathrm{I}} (x_i + y_i) = \sum_{\mathrm{F}\subset\mathrm{I}} x_{\mathrm{F}} y_{\mathrm{I-F}},$$

出现在右端的和是对于 I 的所有子集 F 取的.

事实上, 为了各个和式相乘, 我们以所有可能的方式在每个和式中选择一项, 再把选出的这些项相乘, 最后把对于所有可能的选择所得到的结果相加. 于是和 x_i+y_i 的乘积将是一个和, 它的每一项由出现在给定的和式中的某几个和式中的 x_i 和出现在另外几个和式中的 y_j 相乘得到. 对于一个这样的乘积, 用 F 表示决定选择 x_i 的 i 的值的集合, 那么 I − F 就是决定选择 y_j 的 j 的值的集合, 那么显然所选择的项的乘积就是 $x_{\mathrm{F}}y_{\mathrm{I-F}}$, 把所得的结果相加就得到所陈述的公式.

二项式公式就是刚得到的公式的推论: 取 I 为 n 个元素的集合, 并且对于所有的 $i \in \mathrm{I}$, 选择 $x_i = x, y_i = y$. 一般公式的左端正是 $(x + y)^n$. 在其右端, 显然

$$x_{\mathrm{F}} y_{\mathrm{I-F}} = x^r y^{n-r},$$

其中 r 是 F 的元素的个数. 为了得到二项式公式, 剩下的事情就是考虑到如下的事实: 在一个 n 个元素的集合里, 恰有 $\binom{n}{r}$ 个 r 个元素的子集.

6. 环的同态

给定两个环 K 和 L, 称所有从 K 到 L 内的映射 f 为**从 K 到 L 内的一个同态**, 如果对于任意 $x, y \in \mathrm{K}$, 有

$$f(x + y) = f(x) + f(y), \quad f(xy) = f(x)f(y), \quad f(1) = 1.$$

例 7 从 $\mathbf{Z}$ 到 $\mathbf{Z}/p\mathbf{Z}$ 上的典范映射是环的一个同态 (其实, 环 $\mathbf{Z}/p\mathbf{Z}$ 的结构正是为了这个映射是同态而选择的).

例 8 设 X 是一个集合, K 是一个环, 而 L 是从 X 到 K 内的映射环 (例 3); 那么, 对于所有 $x \in \mathrm{X}$, 从 L 到 K 内的映射

$$f \to f(x)$$

是从 L 到 K 内的一个同态.

环的同态的性质类似于群的同态 (§7, 第 8 小节和 第 9 小节) 的性质. 给定两个环的同态

$$f : \mathrm{K} \to \mathrm{L}, \quad g : \mathrm{L} \to \mathrm{M},$$

复合映射 $g \circ f$ 还是一个同态.

如果一个同态 $f : \mathrm{K} \to \mathrm{L}$ 是双射的, 逆映射还是同态. 这时称 f 是一个**同构**, 我们称两个环 K 和 L 是同构的, 如果存在从 K 到 L 上的一个同构, 关系

K 和 L 是同构的

是环之间的一个等价关系.

设 $f : \mathrm{K} \to \mathrm{L}$ 是一个环同态, 称使得 $f(x) = 0$ 的 $x \in \mathrm{K}$ 的集合为 f 的**核**, 记作

$$\mathrm{Ker}(f)$$

(于是这是当把 f 看作加法群 K 到加法群 L 的同态时 f 的核). 当且仅当 $\mathrm{Ker}(f) = \{0\}$, 同态 f 是单射的.

关系

$$f(x - y) = f(x) - f(y), \quad f(axb) = f(a)f(x)f(b)$$

直接表明环 K 到另一个环内的同态的核 I 满足两个条件:

(i) *I 是加法群 K 的一个子群;*

(ii) *对于任何 $a, b \in \mathrm{K}$ 和 $x \in \mathrm{I}$ 有关系 $axb \in \mathrm{I}$.*

如果环 K 的一个子集 I 满足上面的条件 (i) 和 (ii), 则称为一个**双侧理想**.

注 5 环 K 的一个子集 I 称为 K 的一个**左理想**, 如果 I 是 K 的一个子群, 并且对于任意 $a \in \mathrm{K}$ 和 $x \in \mathrm{I}$ 有关系 $ax \in \mathrm{I}$ (换句话说, 如果所有 $x \in \mathrm{I}$ 的左倍元都在 I 内); 也就是说, I 是 K 的一个非空子集, 并且具有下列性质:

$$ux + vy \in \mathrm{I}, \quad \text{任意 } u, v \in \mathrm{K} \text{ 和 } x, y \in \mathrm{I}.$$

同样定义 K 的一个**右理想**, 这是 K 的一个非空子集, 并且具有下列性质:

$$xu + yv \in \mathrm{I}, \quad \text{任意 } u, v \in \mathrm{K} \text{ 和 } x, y \in \mathrm{I}.$$

双侧理想显然是 K 的子集, 同时是左理想和右理想.

如果 K 是交换的, 左理想、右理想、双侧理想的概念显然是一致的. 这时简单地说理想以代替左理想、右理想、双侧理想.

例 9 如果 K 是一个域, 那么 K 的仅有的左理想是 $\{0\}$ 和 K 自身. 事实上, 如果 K 的左理想 I 含有非零元 a, 那么 I 也含有 $a^{-1}a = 1$, 从而对于任何 $u \in \mathrm{K}$, 含有 $u \cdot 1 = u$, 故得 $\mathrm{I} = \mathrm{K}$.

读者作为习题证明当 $1 \neq 0$ 时这个性质刻画了环成为域的特征.

例 10 设 K 是一个交换环, 对于所有 $x \in \mathrm{K}$, 用 $x\mathrm{K}$ 表示 x **在 K 内的倍元**的集合, 所谓 x 在 K 内的倍元就是形如 ux 的元素, 其中 $u \in \mathrm{K}$. 这个集合 $x\mathrm{K}$ 是 K 的一个理想 (这种类型的理想称为 K 的**主理想**).

称其所有理想都是主理想的交换整环为**主理想整环**. 环 $\mathbf{Z}$ 是一个主理想整环 ($\mathbf{Z}$ 的一个理想是 $\mathbf{Z}$ 的一个子群, 根据 §7 的例 8, 其形式是 $n\mathbf{Z}$). 在 §31 将看到 (如果愿意, 读者现在就可以学习) 有理整数的整除性可以推广到一个主理想整环的元素.

存在不是主理想整环的整环, 最简单的例子是形如

$$x + y\sqrt{10}, \quad 其中\ x, y \in \mathbf{Z}$$

的数组成的 $\mathbf{R}$ 的子环. 这个环和与之类似但更复杂的环的引进引导 19 世纪的数学家 —— 居于首位的是 Dedekind—— 创立了理想这个概念, 后来, 它显示出在数学的许多其他分支里是必需的.

§8 习题

1. 设 K 是一个环 (不假定是交换的).

a) 证明如果 K 的两个元素 x 和 y 是交换的 (即 $xy = yx$), 则对于所有整数 $n \geqslant 1$ 有

$$x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1}).$$

b) 称 K 的一个元素 x 是幂零的, 如果存在一个整数 $n \geqslant 1$, 使得

$$x^n = 0.$$

证明这时 $1 - x$ 在 K 中是可逆的.

c) 如果 K 的两个幂零的元素 x 和 y 是交换的, 则 $x + y$ 是幂零的 (利用对于一个适当大的指数的二项式公式), xy 也是幂零的.

d) 设 a 是 K 的一个元素, 考虑由

$$u(x) = ax - xa \quad 对于所有\ x \in \mathrm{K}$$

定义的从 K 到 K 内的一个映射 u. 说明如果 $a^2 = 0$ 则对于所有 $x \in \mathrm{K}$ 有 $u^3(x) = 0$, 并且如果 $a^3 = 0$ 则对于所有 $x \in \mathrm{K}$ 有 $u^5(x) = 0$. 用一般的方式说明, 如果 a 是幂零的, 则存在一个整数 q, 使得

$$u^q(x) = 0 \quad 对于所有\ x \in \mathrm{K}.$$

说明有

$$u^p(x) = \sum_{k=0}^{p} (-1)^k \binom{p}{k} a^{p-k} x a^k.$$

e) 称 K 的一个元素 u 是幂幺的, 如果 $1 - u$ 是幂零的. 证明如果 $u, v \in \mathrm{K}$ 是幂幺且交换的元素, 则 uv 也是幂幺的. 说明 K 的所有幂幺的元素是可逆的, 并且其逆是幂幺的元素.

[对于一个环的幂幺和幂零的元素的例子, 参见 §14 的习题 10 以及 §19 的习题 19. 在分析中, 有限展开同样提供幂幺元素的例子: 考虑定义在 $t=0$ 的一个邻域内的实变量 t 的函数 $f(t)$ 的环和一个选定的整数 $n \geqslant 1$, 过渡到关于当 t 趋于 0 时的 $o(t^n)$ 类函数的理想的商 (见下面的习题 7, c)); 商环显然有非零幂零元素, 例如函数 t 在这个商环内的像.]

¶¶2. 设 K 是一个环, 假定有理数域 $\mathbf{Q}$ 是 K 的一个子环 (这就允许用有理数乘所有的 $x \in \mathrm{K}$, 和特别地用非零有理整数除所有的 $x \in \mathrm{K}$).

a) 设 x 是 K 的一个幂零元; 定义(*)

$$\exp(x) = 1 + \frac{x}{1} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots,$$

借助二项式公式说明, 如果 $x, y \in \mathrm{K}$ 是幂零的和交换的, 则有

$$\exp(x+y) = \exp(x)\exp(y).$$

b) 设 u 是 K 的一个幂幺元 (习题 1); 定义

$$\log u = -\frac{1-u}{1} - \frac{(1-u)^2}{2} - \cdots - \frac{(1-u)^n}{n} - \cdots,$$

说明如果 $u, v \in \mathrm{K}$ 是幂幺的和交换的, 则有

$$\log(uv) = \log u + \log v.$$

c) 设 x 是 K 的一个幂零元. 证明 $\exp(x)$ 是幂幺的, 并且有

$$\log(\exp(x)) = x.$$

d) 设 x 是 K 的一个幂幺元. 证明 $\log x$ 是幂零的, 并且有

$$\exp(\log u) = u.$$

e) 对于 K 的所有幂零元, 定义

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \cdots,$$
$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \cdots.$$

(*) 这些定义显然是根据分析中讨论的函数

$$e^t, \quad \log(1+t), \quad \cos t \text{ 和 } \sin t$$

的幂级数展开. 这里没有任何收敛性问题, 因为 "级数" 实际上是有限和. 习题在于按照纯粹代数的方案变换它们之间存在的关系, 例如熟知的通常指数函数的性质

$$e^{x+y} = e^x e^y$$

和该函数的幂级数展开的性质. 经常会遇到这样的情形, 即发现涉及分析概念即极限过程的现象存在纯代数的类似.

证明如果 $x, y \in \mathrm{K}$ 是幂零的和交换的, 则有

$$\cos(x+y) = \cos x \cdot \cos y - \sin x \cdot \sin y,$$
$$\sin(x+y) = \sin x \cdot \cos y + \sin y \cdot \cos x.$$

证明对于 K 的所有幂零元 $x \in \mathrm{K}$ 有

$$\cos^2 x + \sin^2 x = 1;$$

当然这里令 $\cos^2 x = \cos x \cdot \cos x$ 和 $\sin^2 x = \sin x \cdot \sin x$.

3. 设 K 是一个环, 对于任意 $x, y \in \mathrm{K}$, 令

$$[x, y] = xy - yx.$$

证明 Jacobi **等式**

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0.$$

4. 在一个环 K 内, 考虑满足关系

$$[h, x] = 2x, \quad [h, y] = -2y, \quad [x, y] = h$$

的元素 x, y, h.

a) 证明公式

$$[h, x^n] = 2n \cdot x^n, \quad [h, y^n] = -2n \cdot y^n.$$

b) 证明 K 的元素

$$4xy + h^2 - 2h$$

关于 x, y 和 h 是交换的.

c) 证明含有 x, y 和 h 的 K 的最小子环是这样的元素的集合, 它们可以写成形式为

$$ax^i y^j h^k$$

的有限项的和, 其中 i, j, k 是自然数, 而 a 是有理整数.

¶5. 设 p 是一个素数. 用 $\mathbf{Z}_p$ 表示可以写成其分母不被 p 整除的分数的形式的 $x \in \mathbf{Q}$ 的集合.

a) 证明 $\mathbf{Z}_p$ 是 $\mathbf{Q}$ 的一个子环.

b) 对于所有 $x \in \mathbf{Q}$, 或者 $x \in \mathbf{Z}_p$, 或者 $x^{-1} \in \mathbf{Z}_p$.

c) $\mathbf{Q}$ 的仅有的包含 $\mathbf{Z}_p$ 的子环是 $\mathbf{Z}_p$ 和 $\mathbf{Q}$.

d) 对于环 $\mathbf{Z}_p$ 的所有理想 I, 存在唯一的一个整数 $n \geqslant 0$, 使得 I 是由 p^n 生成的 (即由 $p^n u$ 组成, 其中的 $u \in \mathbf{Z}_p$).

e) 对于所有非零的 $x \in \mathbf{Q}$, 存在唯一的 $n \in \mathbf{Z}$, 使得

$$x = p^n \cdot u,$$

其中的 u 是环 $\mathbf{Z}_p$ 的可逆元.

f) 对于所有 $x \in \mathbf{Q}$, 令 $v_p(x) = n$, 这里 n 是 e) 中的 n; 此外定义(*)

$$v_p(0) = +\infty.$$

证明对于任意 $x, y \in \mathbf{Q}$ 有

$$v_p(xy) = v_p(x) + v_p(y),$$
$$v_p(x+y) \geqslant \min(v_p(x), v_p(y)),$$

并且 $\mathbf{Z}_p$ 是使得 $v_p(x) \geqslant 0$ 的 $x \in \mathbf{Q}$ 的集合.

g) 证明对应于所有素数 p 的 $\mathbf{Q}$ 的子环 $\mathbf{Z}_p$ 的交集是有理整数环 $\mathbf{Z}$.

¶6. 设 K 是一个域, 而 A 是 K 的一个子环. 称 A 是 K 的一个**赋值环**, 如果 $\mathrm{A} \neq \mathrm{K}$, 并且有

$$x \in \mathrm{A} \text{ 或 } x^{-1} \in \mathrm{A} \quad \text{对于所有非零 } x \in \mathrm{A}.$$

说明环 A 的非可逆元组成 A 的一个理想 $\mathfrak{m}$, 并且异于整个 A 的 A 的所有理想都包含于 $\mathfrak{m}$ 内 [所以在下面习题 7, d) 的术语的意义下 $\mathfrak{m}$ 是 A 的唯一的极大理想].

称 K 上定义的所有函数 v 为 K 的**离散赋值**, 如果 v 的值是有理整数或 $+\infty$, 并且还具有下列性质:

$$v(0) = +\infty; \quad \text{如果 } x \neq 0 \text{ 则 } v(x) \in \mathbf{Z};$$
$$v(xy) = v(x) + v(y), \quad \text{任意 } x, y \in \mathrm{K};$$
$$v(x+y) \geqslant \min(v(x), v(y)), \quad \text{任意 } x, y \in \mathrm{K}.$$

假定 v 是非平凡的 (即不缩减为 0 和 $+\infty$). 说明使得 $v(x) \geqslant 0$ 的 $x \in \mathrm{K}$ 的集合 A 是 K 的一个赋值环, 并且 A 的极大理想 $\mathfrak{m}$ 是使得 $v(x) > 0$ 的 $x \in \mathrm{K}$ 的集合. 选择一个元素 $\pi \in \mathfrak{m}$, 使得 $v(\pi)$ 是最小的; 说明 $\mathfrak{m} = \mathrm{A}\pi$, 并且 A 的所有理想对于一个整数 $n \geqslant 0$ 有形式 $\mathrm{A}\pi^n$.

证明域 $\mathbf{Q}$ 的仅有的赋值环是习题 5 的环 $\mathbf{Z}_p$. 找出 $\mathbf{Q}$ 的所有离散赋值.

7. 设 I 是环 K 的双侧理想, 把关系 $x - y \in \mathrm{I}$ (模 I **同余**) 记作

$$x \equiv y (\mathrm{mod\ I}).$$

a) 说明这是集合 K 上的一个等价关系. 如果 $\mathrm{K} = \mathbf{Z}$ 且 $\mathrm{I} = p\mathbf{Z}$, 会发生什么?

b) 说明关系

$$x' \equiv y' (\mathrm{mod\ I}) \quad \text{和} \quad x'' \equiv y'' (\mathrm{mod\ I})$$

蕴含

$$x' + x'' \equiv y' + y'' (\mathrm{mod\ I}) \quad \text{和} \quad x'x'' \equiv y'y'' (\mathrm{mod\ I}).$$

(*) 用符号 $+\infty$ 表示这样一个对象, 它遵守下列计算规则, 并且仅这些规则 (即下面没有定义的运算没有任何意义):

$$n + (+\infty) = +\infty \quad \text{对于任何 } n \in \mathbf{Z}; \quad (+\infty) + (+\infty) = +\infty;$$

并且约定

$$+\infty > n \quad \text{对于所有 } n \in \mathbf{Z}; \quad +\infty \geqslant +\infty.$$

由此推得例如 $\max(2, +\infty) = +\infty$. 当然我们定义 v_p 时可以回避符号 $+\infty$: 只需不给 $v_p(0)$ 以意义, 而在宣布要证明的关系时不提及 $v_p(0)$. 这种方法将使情况复杂化.

c) 用 K/I 表示 K 关于所考虑的等价关系的商集, 并且用 θ 表示从 K 到 K/I 上的典范映射. 说明在集合 K/I 上存在唯一的一个环结构, 使得映射 θ 是一个同态 (模仿对于环 $\mathbf{Z}/p\mathbf{Z}$ 所给的结构). 说 K/I 是 K 关于双侧理想 I 的**商环**.

d) 假定 K 是交换的. 称 K 的一个理想 I 是**极大的**, 如果 $\mathrm{I} \neq \mathrm{K}$, 并且包含 I 的 K 的仅有的理想是 I 和 K. 说明 I 是极大的, 必须并且只需商环 K/I 是一个域 (注意一个域不具有任何本身和 $\{0\}$ 之外的理想, 反之亦真). 环 $\mathbf{Z}$ 的极大理想是什么?

e) 交换环 K 的一个理想 I 称为**素的**, 如果 $\mathrm{I} \neq \mathrm{K}$, 并且对于 $x, y \in \mathrm{K}$,

$$\text{关系 } xy \in \mathrm{I} \text{ 蕴含 } x \in \mathrm{I} \text{ 或 } y \in \mathrm{I}.$$

说明这个条件表明商环 K/I 是整环. 环 $\mathbf{Z}$ 的素理想是什么?

f) 说明所有的极大理想是素理想. [注意: 其逆仅对于十分特殊种类的环才成立.]

g) 设 K 是一个域, 而 A 是 K 的一个子环. 假定所有的 $x \in \mathrm{K}$ 可以表示成形式 u/v, 其中的 $u, v \in \mathrm{A}$, 并且 $v \neq 0$ (这表明 K 是 A 的分式域, 见 §29). 设(*) $\mathfrak{p}$ 是 A 的一个素理想, 用 $\mathrm{A}_{\mathfrak{p}}$ ($\mathfrak{p}$ 的局部环) 表示形如

$$u/v, \quad \text{其中 } u, v \in \mathrm{A} \text{ 并且 } v \notin \mathfrak{p}$$

的 K 的元素的集合. 说明 $\mathrm{A}_{\mathfrak{p}}$ 是 K 的一个仅有一个极大理想的子环. 并且如果令环 $\mathrm{A}_{\mathfrak{p}}$ 的每个理想 $\mathrm{I} \neq \mathrm{A}_{\mathfrak{p}}$ 对应它同 A 的交集 $\mathrm{I} \cap \mathrm{A}$, 就定义了一个从 $\mathrm{A}_{\mathfrak{p}}$ 的所有理想 $\mathrm{I} \neq \mathrm{A}_{\mathfrak{p}}$ 的集合到包含 $\mathfrak{p}$ 的 A 的所有理想的集合上的一个双射.

8. 设 A 和 B 是两个环. 说明通过给集合 $\mathrm{A} \times \mathrm{B}$ 配备由公式

$$(x', y') + (x'', y'') = (x' + x'', y' + y''), \quad (x', y') \cdot (x'', y'') = (x'x'', y'y'')$$

定义的运算得到一个环, 称为 A 和 B 的**直积**. **直积**可以是一个整环吗?

¶9. 设 m 和 n 是互素的有理整数.

a) 证明对于任意 $a, b \in \mathbf{Z}$ 存在 $x \in \mathbf{Z}$, 使得

$$x \equiv a (\mathrm{mod}\ m), \quad x \equiv b (\mathrm{mod}\ n),$$

并且 x 模 mn 的类由 a 模 m 的类和 b 模 n 的类完全确定. 例子: 求下列同余组的解:

$$x \equiv 4 (\mathrm{mod}\ 7), \quad x \equiv 9 (\mathrm{mod}\ 11).$$

b) 考虑环 $\mathrm{A} = \mathbf{Z}/m\mathbf{Z}, \mathrm{B} = \mathbf{Z}/n\mathbf{Z}$ 和 $\mathrm{C} = \mathbf{Z}/mn\mathbf{Z}$. 借助上一个问题 a), 构造从直积成 $\mathrm{A} \times \mathrm{B}$ (习题 8) 到环 C 上的一个同构. (可以利用 §4 的定理 3.)

c) 设 $q_1, \cdots, q_h$ 是互素整数. 通过对 h 的归纳说明对于任意 $a_1, \cdots, a_h \in \mathbf{Z}$, 存在 $x \in \mathbf{Z}$, 满足 h 个关系

$$x \equiv a_i (\mathrm{mod}\ q_i) \quad (1 \leqslant i \leqslant h).$$

(这个结果称为**中国剩余定理**, 中国的天文学家用其特殊情况来确定某些与天文或其他现象的周期有关事件的日期.)

(*) 表示理想的传统是利用歌德体; 在 §8 的文本中没有顺从这一传统, 为的是免得打扰初学者.

d) 设

$$n = p_1^{r_1} \cdots p_h^{r_h}$$

是整数 n 的素因数分解. 说明环 $\mathbf{Z}/n\mathbf{Z}$ 同构于 h 个环 $\mathbf{Z}/q_i\mathbf{Z}$ $(1 \leqslant q_i \leqslant h)$ 的直积, 这里令

$$q_i = p_i^{r_i} \quad (1 \leqslant i \leqslant h).$$

e) 设 $q_1, \cdots, q_h$ 是任意有理整数, 同余方程组

$$x \equiv a_i (\bmod\ q_i) \quad (1 \leqslant i \leqslant h)$$

有解, 必须并且只需

$$a_i \equiv a_j (\bmod\ d_{ij}) \quad \text{对于 } 1 \leqslant i < j \leqslant h,$$

其中的 d_{ij} 表示 q_i 和 q_j 的最大公约数.

10. 设 I 和 J 是交换环 K 的两个理想. 把可以写成形式 $x + y, x \in \mathrm{I}$ 和 $y \in \mathrm{J}$ 的 K 的元素的集合记作 I + J (理想 I 和 J 的**和**). 把具有下列性质的 $z \in \mathrm{K}$ 的集合记作 IJ (理想 I 和 J 的**积**): 存在一个整数 $n \geqslant 1$, I 的元素 $x_1, \cdots, x_n$ 和 J 的元素 $y_1, \cdots, y_n$, 使得 $z = x_1y_1 + \cdots + x_ny_n$.

a) 说明 I + J 是 K 的包含 I 和 J 的最小理想. 说明 IJ 也是一个包含于 $\mathrm{I} \cap \mathrm{J}$ 内的理想. 建立关系

$$\mathrm{I} + \mathrm{J} = \mathrm{J} + \mathrm{I}, \quad \mathrm{I} + (\mathrm{I}' + \mathrm{I}'') = (\mathrm{I} + \mathrm{I}') + \mathrm{I}'',$$
$$\mathrm{IJ} = \mathrm{JI}, \quad \mathrm{I}(\mathrm{I}'\mathrm{I}'') = (\mathrm{II}')\mathrm{I}'', \quad \mathrm{I}(\mathrm{J}' + \mathrm{J}'') = \mathrm{IJ}' + \mathrm{IJ}''.$$

其中 $\mathrm{I}, \mathrm{I}', \cdots$ 表示 K 的理想. 当 I 和 J 是主理想时解释 I + J 和 IJ.

b) 称 K 的两个理想 I 和 J 是互素的, 如果 I + J = K. 当 $\mathrm{K} = \mathbf{Z}$ 时这个性质的意义是什么? 说明如果 I 和 J 是互素的, 则 $\mathrm{I} \cap \mathrm{J} = \mathrm{IJ}$.

c) 理想 I 和 J 是互素的, 必须并且只需对于任意 $a, b \in \mathrm{K}$, 存在 $x \in \mathrm{K}$, 使得

$$x \equiv a (\bmod\ \mathrm{I}) \quad \text{和} \quad x \equiv b (\bmod\ \mathrm{J}).$$

由此推出商环 K/IJ (习题 7) 同构于环 K/I 和 K/J 的**直积**.

d) 设 $\mathrm{I}, \mathrm{J}_1, \cdots, \mathrm{J}_r$ 是 K 的理想, 假定对于 $1 \leqslant k \leqslant r$, I 和 J_k 是互素的. 说明 I 和积 $\mathrm{J}_1 \cdots \mathrm{J}_r$ 是互素的.

e) 设 $\mathrm{J}_1, \cdots, \mathrm{J}_r$ 是两两互素的. 证明

$$\mathrm{J}_1 \cdots \mathrm{J}_r = \mathrm{J}_1 \cap \cdots \cap \mathrm{J}_r.$$

f) 设 $\mathrm{J}_1, \cdots, \mathrm{J}_r$ 是两两互素的. 说明对于任意 $a_1, \cdots, a_r \in \mathrm{K}$ 可以找到一个 $x \in \mathrm{K}$, 使得

$$x \equiv a_k (\bmod\ \mathrm{J}_k) \quad \text{对于 } 1 \leqslant k \leqslant r.$$

g) K 的两个极大理想 (习题 7, d)) 是互素的, 只要它们是不同的.

¶11. 设 $\mathrm{I}_1, \cdots, \mathrm{I}_r$ 是一个交换环 K 的理想. 如果 K 的一个素理想 (习题 7, e)) 包含积 $\mathrm{I}_1 \cdots \mathrm{I}_r$, 则它至少包含 $\mathrm{I}_1, \cdots, \mathrm{I}_r$ 中的一个.

设 I 是 K 的一个非素的理想. 说明存在 K 的具有下列性质的理想 J′ 和 J″: J′ 和 J″ 包含 I 但是异于 I, 并且 I 包含积理想 J′J″.

12. 给定交换环 K 的一个理想, 称对于一个整数 $n \geqslant 1$ 有

$$x^n \in \mathrm{I}$$

的 $x \in \mathrm{K}$ 的集合为 I 的**根**. 在这个习题里, 用记号

$$\sqrt{\mathrm{I}}$$

表示一个理想 I 的根 (正如下面的公式所指出的, 这个记号是极其不适当的).

a) 证明一个理想 I 的根还是一个理想. 如果 $\mathrm{I} = \{0\}$, 会发生什么? K 的素理想 (习题 7, e)) 的根是什么?

b) 证明下列公式, 其中的 I 和 J 表示环 K 的任意两个环:

$$\sqrt{\mathrm{IJ}} = \sqrt{\mathrm{I} \cap \mathrm{J}} = \sqrt{\mathrm{I}} \cap \sqrt{\mathrm{J}},$$
$$\sqrt{\mathrm{I} + \mathrm{J}} = \sqrt{\sqrt{\mathrm{I}} + \sqrt{\mathrm{J}}},$$
$$\sqrt{\sqrt{\mathrm{I}}} = \sqrt{\mathrm{I}}.$$

c) 完整地确定有理整数环 $\mathbf{Z}$ 的一个理想的根.

13. 称交换环 K 的一个理想是**准素的**, 如果 $\mathrm{I} \neq \mathrm{K}$, 并且对于 K 的任意元素 x, y, 如果有

$$xy \in \mathrm{I}, \quad \mathrm{X} \notin \mathrm{I},$$

那么存在一个整数 $n \geqslant 1$, 使得

$$y^n \in \mathrm{I}.$$

证明一个准素理想的根是一个素环.

¶I (假定异于 K) 是准素的, 必须并且只需在商环 K/I 内, 所有的零因子是幂零的. 环 $\mathbf{Z}$ 的准素理想是什么?

14. 设 $\mathfrak{m}$ 是交换环 K 的一个极大理想. 说明 $\mathfrak{m}$ 的幂

$$\mathfrak{m}^n = \mathfrak{m} \cdots \mathfrak{m} \ (n \text{ 个因子})$$

是准素理想, 其根是 $\mathfrak{m}$.

15. 设 $\mathfrak{m}$ 是交换环 K 的一个极大理想, 而 $\mathfrak{a}$ 是 K 的一个包含于 $\mathfrak{m}$ 内的理想. 假定 $\mathfrak{m}$ 的每一个元素具有一个在 $\mathfrak{a}$ 内的幂. 说明 $\mathfrak{a}$ 是准素的, 并且它的根是 $\mathfrak{m}$.

16. 交换环 K 的一个元素是可逆的, 必须并且只需它不属于任何异于 K 自身的 K 的理想.

我们承认 **Krull 定理**: 给定一个交换环(*) K, 异于 K 的 K 的所有理想包含于 K 的一个极大理想内 [我们提醒, 习题 7, d), K 的一个理想 I 是极大的, 如果 $I \neq K$, 并且包含 I 的仅有的理想是 I 和 K].

证明 K 的一个元素是可逆的, 必须并且只需它不属于 K 的任何极大理想.

17. 设 I 是交换环 K 的所有极大理想的交集. 说明 K 的一个元素 a 属于 I, 当且仅当对于所有 $x \in K$ 元素 $1 + ax$ 是可逆的 (利用习题 16).

[这个结果 (Jacobson) 可以推广到非交换环: 在一个这样的环内极大左理想的交集等于极大右理想的交集; 而这个交集的元素 a 以下列事实为特征: 对于任何 $x, y \in K$, $1 + xay$ 是可逆的. 这个结果的证明完全是初等的.]

¶18. 用 $\mathbf{F}_p$ 表示对于素数 p 的域 $\mathbf{Z}/p\mathbf{Z}$. 给 $\mathbf{F}_{11} \times \mathbf{F}_{11}$ 配备由下列公式给定的运算:

$$(u, v) + (x, y) = (u + x, v + y),$$
$$(u, v) \cdot (x, y) = (ux + 7vy, uy + vx)$$

(其中的 7 显然表示自然数 7 模 11 的类). 证明以这种方式我们得到了一个 121 个元素的域.

19. 证明如果 p 是一个素数, 对于 $1 \leqslant p \leqslant n - 1$, 二项式系数 $\binom{n}{p}$ 是 p 的倍数. 由此推出如果 p 是一个奇素数, 则对于所有整数 $k \geqslant 0$ 有

$$(1 + p)^{p^k} \equiv 1 + p^{k+1} (\text{mod } p^{k+2})$$

(用归纳法证明).

20. 令 $r = \sqrt[3]{2}$. 说明形如

$$a + br + cr^2$$

的数的集合是实数域 $\mathbf{R}$ 的一个子域, 其中的 a, b, c 是任意有理数.

(这个习题对于初学者, 还提供了一个机会来证明 2 不是一个有理数的立方.)

(*) Krull 定理实际上以下列方式推广到所有的环 K, 不论交换与否. 在一个这样的环内, 一个左 (对应的, 右、双侧) 理想 I 称为极大的, 如果 $I \neq K$, 并且包含 I 的 K 的仅有的左 (对应的, 右、双侧) 理想是 I 和 K. 有了这个定义, K 的所有异于 K 的左 (对应的, 右、双侧) 理想包含于 K 的一个极大左 (对应的, 右、双侧) 理想内.

当 $K = \mathbf{Z}$ 时, Krull 定理的含义是所有整数 $n \geqslant 2$ 至少具有一个素因数. 因此 Krull 定理可以看作这个结果对于**所有环的一个没有例外的推广**. 基于这个理由, 尽管其陈述如此简单, 却是整个代数学的最有用的结果之一. 有人 (Gelfand) 甚至在 20 年前把它用在分析的困难定理的证明当中, 其中的环的元素是一个或多个实变量的满足某些条件的函数.

Krull 定理的一般证明是容易的, 只要充分懂得集合论和超限数论 (这是超限数 "大" 理论证明的一个 "具体的" 而又十分不显然的结果的最重要的实际应用之一). 在 §18 将会看到对于 Noether 环 (但是在分析中研究的鲜有这类环) Krull 定理的一个初等证明. 一般证明的思路是这样的, 如果环 K 不包含任何极大理想, 那么就可以在 K 内构造一个无限的甚至是 "超限" 增加的理想的链 (即每一个基数 α 对应一个理想 I_α, 使得对于 $\alpha < \beta$, 理想 I_α 严格包含于 I_β 内 —— 对于有限基数 α 构造显然是容易的, 所有的困难在于推广关于自然数的归纳法到更大范围). 这将与下列事实矛盾: 不存在所有基数的集合 (按通行说法) 到给定集合 (当前是 K 的理想的集合) 的一个单射, 因为不存在所有基数的集合.

§9 复数

1. 平方根

设 K 是一个交换环. 称 K 的一个元素 d 是 K 内的一个**平方**, 如果存在一个 $x \in K$ 使得

$$x^2 = d;$$

这时就称 x 是 K 内 d 的一个**平方根**. 显然如果 x 是 K 内 d 的一个平方根, 则 $-x$ 也是. 如果进而环 K 是整环, d 不可能有多于两个的 K 内的平方根, 因为关系 $x^2 = y^2$ 在所有情形下都可以写成形式 $(x-y)(x+y) = 0$, 当 K 是整环时, 这蕴含 $y = x$ 或 $y = -x$.

自然可能会碰到交换环 K 的一个元素 d 不是 K 内的一个平方: 如果 K 是实数域 $\mathbf{R}$, 当且仅当 $d \geqslant 0$, d 是 K 内的一个平方. 在有理数域 $\mathbf{Q}$ 内, 2 不是一个平方 (然而在实数域 $\mathbf{R}$ 内 2 是一个平方).

这就引导我们考察下列问题: 设 K 是一个交换环, 而 d 是 K 的这样的元素, 它不是 K 内的平方, 是否可以构造一个具有下列性质的交换环 L: K 是 L 的一个子环, 并且 d 是 L 内的一个平方.

当 $K = \mathbf{R}$ 和 $d = -1$ 时这个问题的解决历史性地导致 "复数" 的发明, 在本节的后面我们将定义复数. 为了真正懂得这些 "数" 的构造, 研究我们刚提出的一般问题是必需的 (却不是更困难的).

2. 预备知识

设 d 是交换环 K 的一个元素. 在这一小节, 我们假定前一小节陈述的问题已经解决, 并且用 L 表示一个交换环, K 是它的一个子环, ω 是 d 在 L 内的一个平方根.

用 L′ 表示 L 的具有下列性质的元素 z 的集合: 存在 $x, y \in K$ 使得

$$z = x + \omega y. \tag{1}$$

那么 L′ 是 L 的包含 K 的一个子环, 并且在 L′ 中 d 具有一个平方根.

显然 L′ 包含 K (在前面的关系中令 $y = 0$) 以及 ω (令 $x = 0$ 和 $y = 1$), 剩下要做的是看出 L′ 是 L 的一个子环. 由于 L′ 包含 K, 自然含有 -1, 为此只需指出如果 L′ 含有两个元素, 则也含有它们的和与乘积. 而这立刻从以下公式得到:

$$(x' + \omega y') + (x'' + \omega y'') = (x' + x'') + \omega(y' + y''), \tag{2}$$

$$(x' + \omega y') \cdot (x'' + \omega y'') = (x'x'' + dy'y'') + \omega(x'y'' + x''y'), \tag{3}$$

其中的 (3) 由于

$$\omega^2 = d \tag{4}$$

是显然的.

这个结果表明, 如果提出的问题有一个解, 或者说如果构造了 K 的一个交换的母环 L 和一个满足 (4) 的 ω, 那么甚至可以构造 L 使得 L 的所有元素具有形式 (1), 其中 $x, y \in \mathrm{K}$. 换句话说, 如果引进由

$$f(x,y) = x + \omega y$$

给定的映射 $f: \mathrm{K} \times \mathrm{K} \to \mathrm{L}$, 可以假定 f 是满射的.



注 1　如果 K 是一个域, 且 d 不是 K 内的平方, 那么 f 还是单射的. 事实上, 关系 $x' + \omega y' = x'' + \omega y''$ 可以改写为

$$x + \omega y = 0, \quad \text{其中 } x = x' - x'', y = y' - y'';$$

如果 $y \neq 0$, 那么 y 在 K 内是可逆的 (因为 K 是一个域), 当然在 L 内是可逆的, 关系 $x + \omega y = 0$ 蕴含

$$\omega = -y^{-1}x \in \mathrm{K},$$

这跟 d 不是 K 内的平方相矛盾. 于是有 $y = 0$, 因此 $x = 0$, 即 $x' = x'', y' = y''$, 这就证明了 f 是单射的.

引进映射 f, 我们将会看到公式 (2) 和 (3) 就写成

$$f(x', y') + f(x'', y'') = f(x' + x'', y' + y'') \tag{2'}$$

$$f(x', y') \cdot f(x'', y'') = f(x'x'' + dy'y'', x'y'' + x''y'). \tag{3'}$$

这些 (假定问题已经解决所得到的) 公式现在我们要用来作为实际构造所提出的问题的解的出发点.

3. 环 $\mathrm{K}[\sqrt{d}]$

令 d 是交换环 K 的一个元素. 我们要构造一个新的环 L, 按照传统记作

$$\mathrm{K}[\sqrt{d}],$$

其定义如下: 集合 L 是笛卡儿乘积 $\mathrm{K} \times \mathrm{K}$, 其元素是 K 的元素的序偶 (x, y); 在 L 内的两个基本运算由以下公式定义

$$(x', y') + (x'', y'') = (x' + x'', y' + y'') \tag{2'''}$$

$$(x', y') \cdot (x'', y'') = (x'x'' + dy'y'', x'y'' + x''y'), \tag{3'''}$$

公式涉及给定元素 d 和环 K 内的运算.

当然公式 (2‴) 和 (3‴) 使得集合 $\mathrm{K}\times\mathrm{K}$ 成为一个交换环这件事一点儿也不显然, 我们应当予以证明 (正如在 $\mathrm{K}=\mathbf{R}$ 和 $d=-1$ 的古典情形; 在这一特殊情形的证明比一般情形既不更容易, 也不更复杂).

首先, 集合 $\mathrm{K}\times\mathrm{K}$ 配备了 (2‴) 定义的加法是一个交换群: 这由 §7 第 2 小节 (群的直积) 和配备了给定的加法的集合 K 是交换群这一事实得到.

现在指出 (3‴) 定义的乘法是结合的. 事实上, 根据定义, 一方面

$$
\begin{aligned}
&(x,y)[(x',y')(x'',y'')]=(x,y)(x'x''+dy'y'',x'y''+x''y')\\
&=(x(x'x''+dy'y'')+dy(x'y''+x''y'),x(x'y''+x''y')+y(x'x''+dy'y'')),
\end{aligned}
$$

而另一方面

$$
\begin{aligned}
&[(x,y)(x',y')](x'',y'')=(xx'+dyy',xy'+x'y)(x'',y'')\\
&=((xx'+dyy')x''+d(xy'+x'y)y'',(xx'+dyy')y''+x''(xy'+x'y));
\end{aligned}
$$

比较所得到的结果 (当然要利用 K 内的计算规则) 就平凡地得到结合性. 经过类似的计算可以证明交换性和乘法对于加法的分配性.

最后, 关系

$$(1,0)(x,y)=(x,y)$$

表明集合 $\mathrm{K}\times\mathrm{K}$ 配备了 (3‴) 定义的运算具有一个中性元.

于是集合 $\mathrm{K}\times\mathrm{K}$ 配备了 (2‴) 和 (3‴) 定义的运算就是一个交换环.

现在指出 $\mathrm{K}\times\mathrm{K}$ 包含一个同构于 K 的子环. 为此考虑映射 $j:\mathrm{K}\to\mathrm{K}\times\mathrm{K}$, 其定义是

$$j(x)=(x,0),$$

这显然是一个单射. 其次, 一个容易的计算指出

$$j(x')+j(x'')=j(x'+x''),\quad j(x')j(x'')=j(x'x''),\quad j(-1)=(-1,0)=-1,$$

其中, 最后一个关系的右端的 -1 表示环 $\mathrm{K}\times\mathrm{K}$ 的单位元 $1=(1,0)$ 的相反元. 前面这些公式显然指出 j 是从 K 到 $\mathrm{K}\times\mathrm{K}$ 的一个子环上的同构.

既然 j 变换 K 上的运算成 $\mathrm{K}\times\mathrm{K}$ 的子环 $j(\mathrm{K})$ 上的运算, 把 K 的每个元素 x 与 $\mathrm{K}\times\mathrm{K}$ 的元素等同就没有任何不便, 今后我们就这么办(*).

(*) 这里用的方法跟把有理整数跟特殊的有理数等同 (或把有理数跟特殊的实数等同) 类似. 不管用以数学地定义有理整数和有理数的过程是什么, 有理整数 n 跟用 $\frac{n}{1}$ 表示的有理数从来就表示真的相等; 可以简单地说映射 $n\to\frac{n}{1}$ 是单射的, 并且把关于有理整数的代数运算变换为对应有理数的代数运算. 这就解释了为什么在实际中人们不区分整数 n 和有理数 $\frac{n}{1}$.

为了说明我们这样就解决了在第 1 小节所提出的问题, 还需要判断 K 的元素 d 在 $K[\sqrt{d}]$ 里是一个平方. 考虑 L 的元素

$$\omega = (0,1). \tag{5}$$

平凡的计算指出

$$\omega^2 = (0,1)(0,1) = (0\cdot 0 + d\cdot 1\cdot 1, 0\cdot 1 + 1\cdot 0) = (d,0) = j(d),$$

由于我们已经按照一般的方式约定把每一个元素 $x \in K$ 与 L 的元素 $j(x)$ 等同, 我们的断言至此证明完毕.

我们不仅仅解决了第 1 小节所提出的问题, 而且构造了一个典范解 $K[\sqrt{d}]$. 我们说它由 K 添加 d 的一个平方根得到; 形如 $K[\sqrt{d}]$ 的环称为 K 的一个**二次扩张**.

为了结束这个构造过程, 我们给 $K[\sqrt{d}]$ 的元素指定一个方便的记号. 首先所有 $x \in K$ 与 L 的元素 $(x,0)$ 的等同允许写出

$$(x,0) = x \tag{6}$$

(严格说来这其实也是错误的). 直接的计算指出对于任意 $x, y \in K$ 有公式

$$(x,y) = (x,0) + (0,1)(y,0);$$

考虑到 (5) 和 (6), 此式可以写成

$$(x,y) = x + \omega y, \tag{7}$$

而这就回到了第 2 小节中假定问题已经解决所研究的情形.

总之, 环 $K[\sqrt{d}]$ 的所有元素以唯一的一种方式写成形式 $x + \omega y$, 这个事实结合交换环的公理和关系

$$\omega^2 = d$$

足以使我们能进行所需要的所有运算. 换句话说, 现在读者可以忘掉环 $K[\sqrt{d}]$ 的构造过程, 而只记住其性质.

例 1 最重要的情形是实数域 $K = \mathbf{R}$ 和 $d = -1$, 这样得到的环 $\mathbf{R}[\sqrt{-1}]$ 记作 $\mathbf{C}$, 它的元素称为**复数**. 一个复数是实数的一个序偶 (x,y), 在复数上借助对于 $d = -1$ 的公式 $(2''')$ 和 $(3''')$ 进行运算, 即令

$$(x', y') + (x'', y'') = (x' + x'', y' + y''),$$
$$(x', y')\cdot(x'', y'') = (x'x'' - y'y'', x'y'' + x''y').$$

在实际中, 我们不使用这些公式, 我们仅利用复数的下列性质:

a) 复数组成一个交换环 $\mathbf{C}$ (后面会看到 $\mathbf{C}$ 是一个域);

b) 实数域 $\mathbf{R}$ 是 $\mathbf{C}$ 的一个子域;

c) 存在一个复数 i (代替二次扩张中使用的记号 ω) 使得

$$i^2 = -1;$$

d) 所有复数 z 用一种并且仅一种方式写成形式

$$z = x + iy,$$

其中的 x 和 y 是实数.

对于复数的计算再没有什么需要知道的了. 例如

$$\begin{aligned}(2+3i)(5-7i) &= 2\cdot 5 - 2\cdot 7i + 3i\cdot 5 - 3i\cdot 7i\\ &= 10 - 14i + 15i - 21i^2\\ &= 10 + i + 21 = 31 + i.\end{aligned}$$

给定一个复数

$$z = x + iy,$$

其中 x 和 y 是实数, 我们说 x 是 z 的**实部**, 而 y 是 z 的**虚部**, 把它们记作

$$x = \mathrm{Re}(z), \quad y = \mathrm{Im}(z).$$

我们说 z 是**纯虚数**, 如果 $\mathrm{Re}(z) = 0$, 但 $\mathrm{Im}(z) \neq 0$; 反之, 说 z 是实数 (即属于 $\mathbf{C}$ 的子环 $\mathbf{R}$), 如果 $\mathrm{Im}(z) = 0$.

4. 二次扩张的可逆元

设 K 是一个交换环, 而 d 是 K 的一个元素. 考虑前一小节定义的二次扩张

$$\mathrm{L} = \mathrm{K}[\sqrt{d}].$$

给定 L 的一个元素

$$z = x + \omega y \quad (x, y \in \mathrm{K}),$$

称 L 的元素

$$\bar{z} = x - \omega y \tag{8}$$

为 z 的**共轭**, 而 K 的元素

$$\mathrm{N}(z) = \bar{z}\cdot z = (x-\omega y)(x+\omega y) = x^2 - \omega^2 y^2 = x^2 - dy^2 \tag{9}$$

称为 z 的**范数**.

对于 L 的任意元素 z' 和 z'', 我们有关系

$$\overline{z'+z''}=\overline{z'}+\overline{z''}, \tag{10}$$

$$\overline{z'z''}=\overline{z'}\cdot\overline{z''}, \tag{11}$$

这两个公式告诉我们如何计算 L 的元素的和与乘积的共轭. 为了证明关系 (10) 和 (11), 最简单的方法莫过于从公式 (2‴) 和 (3‴) 出发, 并且考察当把 y' 和 y'' 换为它们的相反元而不修改 x' 和 x'' 时会发生什么.

关系 (11) 表示对于任意 z' 和 z'',

$$\mathrm{N}(z'z'')=\mathrm{N}(z')\mathrm{N}(z''). \tag{12}$$

事实上,

$$\mathrm{N}(z'z'')=\overline{z'z''}\cdot z'z''=\overline{z'}\cdot\overline{z''}\cdot z'\cdot z''=\overline{z'}\cdot z'\cdot\overline{z''}\cdot z''=\mathrm{N}(z')\cdot\mathrm{N}(z''),$$

正如我们所陈述的. 我们还注意到

$$\mathrm{N}(1)=1. \tag{13}$$

定理 1 设 K 是一个交换环, d 是 K 的一个元素, 而 z 是环 $\mathrm{K}[\sqrt{d}]$ 的一个元素. z 在 $\mathrm{K}[\sqrt{d}]$ 内是可逆的, 必须且只需 $\mathrm{N}(z)$ 在 K 内是可逆的. 这时有

$$z^{-1}=\mathrm{N}(z)^{-1}\cdot\bar{z}. \tag{14}$$

假定 z 是可逆的, 那么考虑到 (12) 和 (13), 关系 $z^{-1}\cdot z=1$ 给出关系

$$\mathrm{N}(z^{-1})\cdot\mathrm{N}(z)=1,$$

因此 $\mathrm{N}(z)$ 必然在环 K 内是一个可逆元.

反之, 假定 $\mathrm{N}(z)$ 在 K 内是可逆的, 关系

$$\bar{z}z=\mathrm{N}(z)$$

蕴含

$$\mathrm{N}(z)^{-1}\cdot\bar{z}\cdot z=1,$$

此即表明 z 是可逆的, 并且它的逆元由关系 (14) 给出, 定理得证.

5. 交换域的情形

现在我们可以证明下列结果:

定理 2 设 d 是交换域 K 内的一个元素, 下列性质是等价的:

a) 环 $\mathrm{K}[\sqrt{d}]$ 是一个域;

b) d 在 K 内不是一个平方.

为了证明 a) 蕴含 b), 假定存在一个 $x \in \mathrm{K}$ 使得

$$x^2 = d,$$

则有 $x^2 = \omega^2$, 于是 $(x-\omega)(x+\omega) = 0$. 如果 $\mathrm{K}[\sqrt{d}]$ 是一个域, 则是一个整环, 从而有 $\omega = x$ 或 $\omega = -x$, 这是不可能的, 因为对于 $x, y \in \mathrm{K}$, 关系 $x + \omega y = 0$ 蕴含 $x = y = 0$.

现在证明 b) 蕴含 a). 设 $z = x + \omega y$ 是 $\mathrm{K}[\sqrt{d}]$ 的一个非零元, 为了证明它是可逆的, 只需证明 (定理 1) $\mathrm{N}(z)$ 在 K 内是可逆的. 由于 K 是一个域, 这就归结为证明 $\mathrm{N}(z) = 0$ 蕴含 $z = 0$, 换句话说

$$x^2 - dy^2 = 0 \quad \text{蕴含} \quad x = y = 0.$$

如果 $y \neq 0$, K 的元素 y 将是可逆的, 于是

$$d = (y^2)^{-1}x^2 = (y^{-1}x)^2,$$

这与 d 在 K 内不是一个平方的假设矛盾. 于是 $y = 0$, 因此 $x^2 = 0$, 从而 $x = 0$. 这就完成了定理的证明.

例 2 取 $\mathrm{K} = \mathbf{R}$ 和 $d = -1$, 我们发现复数环 $\mathbf{C}$ 事实上是一个交换域. 基于这个理由, $\mathbf{C}$ 是一个复数域. 给定一个非零复数

$$z = x + iy,$$

它的逆元由关系 (14) 给定, 而由于这里

$$\mathrm{N}(z) = x^2 + y^2,$$

故

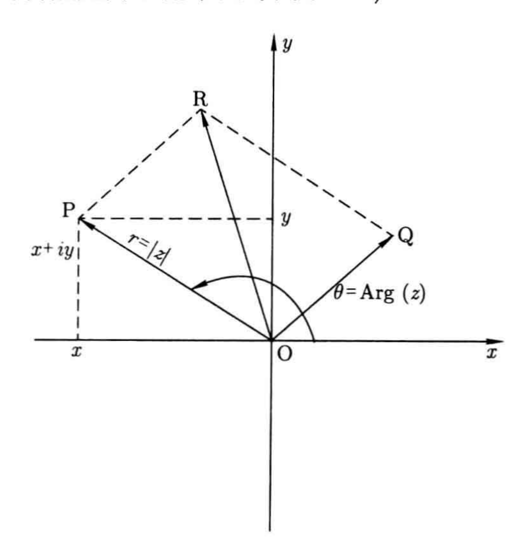
$$z^{-1} = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2},$$

其中的分母仅当 $x = y = 0$ 时即 $z = 0$ 为零 (这与定理 2 的证明符合).

例 3 环 $\mathrm{L} = \mathbf{Q}[\sqrt{2}]$ 事实上是一个交换域. 此外可以把它等同于 $\mathbf{R}$ 的一个子域: 只需令 L 的每个元素 $x + y\omega$ 对应实数 $x + y\sqrt{2}$ (这里 $\sqrt{2}$ 表示数 2 的通常的平方根); 从 L 到 $\mathbf{R}$ 内这样定义的映射是单射的, 并且跟问题中所涉及的运算是相容的. 这样我们又回到在 §8 例 6 中定义的 $\mathbf{R}$ 的子域.

6. 复数的几何表示

在平面上取定两个坐标轴 Ox 和 Oy, 令所有复数 $x+iy$ 对应平面上坐标为 x, y 的点是自然的. 反之, 令平面上所有的坐标为 x, y 的点 P 对应复数 $z=x+iy$, 传统上称为点 P 的**附标**. 这样就得到复数集合 **C** 到平面上点的集合的一个双射 (而此前坐标轴的选取的目的自然是为了把平面等同于 $\mathbf{R}^2$).



复数的 "几何表示" 允许容易地解释复数的加法: 如果 P 和 Q 是平面上的**附标**为 u 和 v 的点, 那么附标为 $u+v$ 的点 R 由关系

$$\overrightarrow{\mathrm{OR}}=\overrightarrow{\mathrm{OP}}+\overrightarrow{\mathrm{OQ}}$$

给定, 因为为了对于起点为 O 的向量求和只需对于它们关于坐标轴 Ox 和 Oy 的分量求和.

设 P 是附标为 $z=x+iy$ 的点, 实数

$$\mathrm{N}(z)=\bar{z}z=x^2+y^2$$

根据 Pythagore (毕达哥拉斯) 定理直观上由

$$\mathrm{N}(z)=\mathrm{OP}^2$$

给定. 从 P 到 O 的距离, 即数

$$r=\mathrm{OP}=\sqrt{x^2+y^2}=\sqrt{\bar{z}z} \tag{15}$$

称为 z 的**模**或**绝对值**, 用记号表示为

$$|z|.$$

我们有 $|z| \geqslant 0$, 并且当且仅当 $z=0$ 时 $|z|=0$. 进而, 如果考察图形, 经典的三角形边长的不等式指出, 对于任意 $u, v \in \mathbf{C}$, 我们有 $|u+v| \leqslant |u|+|v|$, 更一般的, 对于任意复数 $z_1, \cdots, z_n$ 有

$$\boxed{|z_1+\cdots+z_n| \leqslant |z_1|+\cdots+|z_n|.} \tag{16}$$

这个不等式通常称为三角不等式.

假定 $z \neq 0$. 角

$$\theta=(\widehat{\overrightarrow{\mathrm{O}x}, \overrightarrow{\mathrm{OP}}})$$

是一个模 2π 的实数 (§4, 例 10), 称为复数 $z \neq 0$ 的**辐角**, 并且经常用记号

$$\operatorname{Arg}(z)$$

表示. 上面的图形指出 z 的实部和虚部由公式

$$x=r\cos\theta, \quad y=r\sin\theta \tag{17}$$

通过 z 的绝对值 r 和辐角 θ 表示. 于是还可以写出

$$\boxed{z=r(\cos\theta+i\sin\theta),} \tag{18}$$

这个公式称为 z 的**三角表示**. 我们注意, 反之

关系 $z=r(\cos\theta+i\sin\theta)$ (其中 $r>0$) 蕴含 $r=|z|, \quad \theta=\operatorname{Arg}(z)$.

因为后一关系指出 z 的实部和虚部是

$$x=r\cos\theta, \quad y=r\sin\theta,$$

由此得到 $r^2=x^2+y^2$, 由于 r 是正的, 故 $r=|z|$. 记 z 的辐角为 θ', 即得 $\cos\theta=\cos\theta', \sin\theta=\sin\theta'$, 故精确到 2π 的倍数有 $\theta=\theta'$, 这就证明了我们的断言.

我们要从上述结果推出用来计算乘积的模和辐角的公式

$$\boxed{|z_1 z_2|=|z_1|\cdot|z_2|, \quad \operatorname{Arg}(z_1 z_2)=\operatorname{Arg}(z_1)+\operatorname{Arg}(z_2).} \tag{19}$$

为此用模和辐角写出 z_1 和 z_2,

$$z_1=r_1(\cos\theta_1+i\sin\theta_1), \quad z_2=r_2(\cos\theta_2+i\sin\theta_2);$$

就得到

$$z_1 z_2=r_1 r_2(\cos\theta_1+i\sin\theta_1)(\cos\theta_2+i\sin\theta_2),$$

由于 $r_1 r_2$ 是正的, 为了得到所要求的结果, 只需证明关系

$$(\cos\theta_1+i\sin\theta_1)(\cos\theta_2+i\sin\theta_2)=\cos(\theta_1+\theta_2)+i\sin(\theta_1+\theta_2).$$

根据复数的乘法规则, 左端的实部等于

$$\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 = \cos(\theta_1+\theta_2),$$

而虚部等于

$$\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2 = \sin(\theta_1+\theta_2),$$

这就证明了所叙述的结果.

这个结果自然推广到多个因子的乘积, 其形式是

$$\boxed{\prod_{p=1}^{n}(\cos\theta_p + i\sin\theta_p) = \cos\theta + i\sin\theta, \quad \theta = \sum_{p=1}^{n}\theta_p.} \tag{20}$$

特殊情形, 假定 θ_p 都等于同一个角 θ, 就得到 **De Moire 公式**

$$\boxed{(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta.} \tag{21}$$



注 2 公式

$$|z_1z_2| = |z_1|\cdot|z_2|, \quad \operatorname{Arg}(z_1z_2) = \operatorname{Arg}(z_1) + \operatorname{Arg}(z_2)$$

可以给复数乘法一个 "几何" 解释. 设

$$z = r(\cos\theta + i\sin\theta), \quad r \geqslant 0$$

是一个复数, 令平面上所有的点 P 对应点 P′, 其附标是 P 的附标乘以 z 的乘积 (平面上的映射 $\mathrm{P} \to \mathrm{P}'$ 对应从 $\mathbf{C}$ 到 $\mathbf{C}$ 内的映射 $u \to zu$), 那么从 P 到 P′ 的变换是一个相似变换, 精确地说, 是中心为 O 比例为 r 的位似和绕 O 角为 θ 的旋转的乘积.

同样, 关系

$$|z^{-1}| = |z|^{-1}, \quad \operatorname{Arg}(z^{-1}) = -\operatorname{Arg}(z)$$

(在 (19) 中令 $z_1 = z, z_2 = z^{-1}$ 得到) 表明, 先进行对于中心为 O 半径为 1 的圆周的反演, 再进行关于 Ox 轴的对称, 就从附标为 z 的点过渡到附标为 z^{-1} 的点.

复数运算的这些几何解释在分析中经常用到, 而 Gauss 是第一人. 在 1799 年, 在他给出 d'Alembert-Gauss 定理 (代数基本定理) 的第一个严格证明 (§33, 第 2 小节) 时系统地应用了它.

至于复数的发明则远远早于 Gauss, 这要追溯到 16 世纪的意大利数学家.

7. 三角函数的乘法公式

上面确立的对于所有实数 t 都成立的 De Moire 公式

$$(\cos t + i\sin t)^n = \cos nt + i\sin nt$$

使我们能够通过 $\cos t$ 和 $\sin t$ 计算 $\cos nt$ 和 $\sin nt$. 事实上, 左端可以改写成

$$\begin{aligned}&\cos^n t + \binom{n}{1}\cos^{n-1} t\cdot i\cdot \sin t + \binom{n}{2}\cos^{n-2} t(i\sin t)^2 + \cdots + (i\sin t)^n\\ = &\cos^n t + i\binom{n}{1}\cos^{n-1} t\sin t - \binom{n}{2}\cos^{n-2} t\sin^2 t - i\binom{n}{3}\cos^{n-3} t\sin^3 t\\ &+\binom{n}{4}\cos^{n-4} t\sin^4 t + i\binom{n}{5}\cos^{n-5} t\sin^5 t + \cdots + i^n\sin^n t;\end{aligned}$$

而根据 De Moire 公式, 这个表达式的实部和虚部分别等于 $\cos nt$ 和 $\sin nt$, 故得

$$\cos nt = \cos^n t - \binom{n}{2}\cos^{n-2} t\sin^2 t + \binom{n}{4}\cos^{n-4} t\sin^4 t - \cdots, \tag{22}$$

$$\sin nt = \binom{n}{1}\cos^{n-1} t\sin t - \binom{n}{3}\cos^{n-3} t\sin^3 t + \binom{n}{5}\cos^{n-5} t\sin^5 t - \cdots, \tag{23}$$

这就是要找的公式. 在这些公式中最后的项依赖 n 的奇偶性, 因为 i^n 当 n 是偶数时是实数, 而当 n 是奇数时是虚数. 例如

$$\begin{aligned}\cos 4t &= \cos^4 t - 6\cos^2 t\sin^2 t + \sin^4 t,\\ \cos 5t &= \cos^5 t - 10\cos^3 t\sin^2 t + 10\cos t\sin^4 t.\end{aligned}$$

对于 $\cos nt$ 和所获得的公式自然可以用来计算

$$\tan nt = \frac{\sin nt}{\cos nt} = \frac{\binom{n}{1}\cos^{n-1} t\sin t - \binom{n}{3}\cos^{n-3} t\sin^3 t + \binom{n}{5}\cos^{n-5} t\sin^5 t - \cdots}{\cos^n t - \binom{n}{2}\cos^{n-2} t\sin^2 t + \binom{n}{4}\cos^{n-4} t\sin^4 t - \cdots}, \tag{24}$$

最后的分式的分子和分母除以 $\cos^n t$ 便得

$$\tan nt = \frac{\binom{n}{1}\tan t - \binom{n}{3}\tan^3 t + \binom{n}{5}\tan^5 t + \cdots}{1 - \binom{n}{2}\tan^2 t + \binom{n}{4}\tan^4 t - \cdots}.$$

例如, 我们有

$$\tan 5t = \frac{5\tan t - 10\tan^3 t + \tan^5 t}{1 - 10\tan^2 t + 5\tan^4 t}.$$

现在要利用在任何域上只要 $q \neq 1$ 就有效的公式

$$1+q+q^2+\cdots+q^n=\frac{1-q^{n+1}}{1-q},$$

在其中取

$$q=r^2,\quad r=\cos t+i\sin t$$

($t \not\equiv 0(\bmod\ \pi)$), 那么有

$$\frac{1-q^{n+1}}{1-q}=\frac{r^{2n+2}-1}{r^2-1}=\frac{r^{2n+1}-r^{-1}}{r-r^{-1}}=r^n\frac{r^{n+1}-r^{-n-1}}{r-r^{-1}},$$

而 De Moire 公式告诉我们

$$\begin{aligned}r^{n+1}&=\cos(n+1)t+i\sin(n+1)t,\\ r^{-1}&=\cos t-i\sin t,\\ r^{-n-1}&=\cos(n+1)t-i\sin(n+1)t.\end{aligned}$$

于是得到

$$\begin{aligned}\frac{1-q^{n+1}}{1-q}&=(\cos nt+i\sin nt)\frac{2i\sin(n+1)t}{2i\sin t}\\ &=(\cos nt+i\sin nt)\frac{\sin(n+1)t}{\sin t};\end{aligned}$$

另一方面, De Moire 公式给出

$$1+q+q^2+\cdots+q^n=\sum_{k=0}^{n}(\cos 2kt+i\sin 2kt);$$

比较这个表达式和前一个表达式的实部和虚部, 我们得到公式

$$\begin{aligned}1+\cos 2t+\cos 4t+\cdots+\cos 2nt&=\cos nt\frac{\sin(n+1)t}{\sin t},\\ \sin 2t+\sin 4t+\cdots+\sin 2nt&=\sin nt\frac{\sin(n+1)t}{\sin t},\end{aligned}$$

我们更愿意写成

$$1+\cos t+\cos 2t+\cdots+\cos nt=\frac{\cos\left(n\dfrac{t}{2}\right)\sin(n+1)\dfrac{t}{2}}{\sin\dfrac{t}{2}},$$

$$\sin t+\sin 2t+\cdots+\sin nt=\frac{\sin\left(n\dfrac{t}{2}\right)\sin(n+1)\dfrac{t}{2}}{\sin\dfrac{t}{2}}.$$

自然在这些公式内要假定 $\sin\dfrac{t}{2}$ 不是零, 即 t 不是 2π 的整数倍.

§9 习题

1. a) 证明所有复数 $d \neq 0$ 在域 $\mathbf{C}$ 恰好具有两个平方根 (通过 d 的模和辐角求平方根的模和辐角).

b) 由此推出带复系数的二次方程

$$az^2 + bz + c = 0$$

在 $\mathbf{C}$ 至少有一个根, 并且如果

$$b^2 - 4ac \neq 0,$$

则具有两个根.

c) 在 C 内解方程

$$z^2 + (5 - 2i)z + 5 - 5i = 0,$$
$$z^2 + (1 - 2i)z - 2i = 0,$$
$$z^4 - 30z^2 + 289 = 0.$$

2. 求下列复数的实部和虚部:

$$\frac{(1+2i)^2 - (1-i)^3}{(3+2i)^3 - (2+i)^2}; \quad \frac{(1+i)^9}{(1-i)^7}; \quad \left(-\frac{1}{2} + \frac{i\sqrt{3}}{2}\right)^3; \quad \sqrt[4]{2 - i\sqrt{12}};$$
$$(az^2 + bz)(bz^2 + az), \quad \text{其中 } z = -\frac{1}{2} + \frac{i\sqrt{3}}{2}.$$

3. 计算下列复数的模和辐角:

$$-1 + i; \quad -1 - i\sqrt{3}; \quad 2 + \sqrt{3} + i; \quad \left(\frac{1 + i\sqrt{3}}{1 - i}\right)^{20};$$
$$(1 + \cos\theta + i\sin\theta)^n.$$

4. 求下列方程的复数解 (利用 de Moivre 公式):

$$z^3 + 1 = 0; \quad z^5 = 1; \quad z^6 + 27 = 0; \quad z^8 = \frac{1+i}{\sqrt{3} - i}.$$

5. a) 根据角 t 的三角函数计算

$$\cos 5t, \quad \cos 8t, \quad \sin 6t, \quad \sin 9t, \quad \tan 6t.$$

b) 根据角 t 的整倍数的三角函数计算

$$\sin^3 t, \quad \sin^4 t, \quad \cos^5 t, \quad \cos^6 t.$$

6. 设 u 和 v 是两个复数. 证明有关系

$$|u + v|^2 + |u - v|^2 = 2(|u|^2 + |v|^2);$$

其几何解释是什么?

7. 证明下列关系:

$$2^{2n}\cos^{2n}t = 2\cos 2nt + 2\binom{2n}{1}\cos(2n-2)t + \cdots + 2\binom{2n}{n-1}\cos 2t + \binom{2n}{n},$$

$$2^{2n}\cos^{2n+1}t = \cos(2n+1)t + \binom{2n+1}{1}\cos(2n-1)t + \cdots + \binom{2n+1}{n}\cos t,$$

$$2^{2n}\sin^{2n}t = 2\sum_{k=0}^{n}(-1)^{n+k}\binom{2n}{k}\cos 2(n-k)t + \binom{2n}{n},$$

$$2^{2n}\sin^{2n+1}t = \sum_{k=0}^{n}(-1)^{n+k}\binom{2n+1}{k}\sin(2n-2k+1)t,$$

$$\binom{n}{1} - \frac{1}{3}\binom{n}{3} + \frac{1}{9}\binom{n}{5} - \frac{1}{27}\binom{n}{7} + \cdots = \frac{2^n}{3^{\frac{n-1}{2}}}\sin\frac{n\pi}{6},$$

$$\cos\frac{\pi}{11} + \cos\frac{3\pi}{11} + \cos\frac{5\pi}{11} + \cos\frac{7\pi}{11} + \cos\frac{9\pi}{11} = \frac{1}{2}.$$

8. 设 $\mathbf{U}$ 是满足条件 $|z|=1$ 的复数的集合. 指出这是非零复数乘法群 $\mathbf{C}^*$ 的一个子群. 构造群 $\mathbf{U}\times\mathbf{R}_+^*$ 和 $\mathbf{C}^*$ 之间的一个同构. 指出 $\mathbf{U}$ 同构于平面内的绕一个点的旋转群.

9. 称满足条件

$$\mathrm{Im}(z) > 0$$

的复数 z 的集合 P 为 **Poincaré 半平面**, 而满足条件

$$|z| < 1$$

的复数 z 的集合 D 称为**单位圆盘**. 指出

$$z \to \frac{z-i}{z+i}$$

是从 P 到 D 上的一个双射 (用 A 和 B 表示附标为 i 和 $-i$ 的点, 用 M 表示附标为 z 的点, 用三角形 MAB 的角和边长表示 $(z-i)/(z+i)$ 的模和辐角).

10. 用 a, b, c, d 表示满足条件 $ad-bc=1$ 的实数.

a) 对于所有非实数的复数 z 有

$$\mathrm{Im}\left(\frac{az+b}{cz+d}\right) = \frac{\mathrm{Im}(z)}{|cz+d|^2}.$$

b) 设 P 是 Poincaré 半平面 (习题 9), 指出存在 P 的一个置换 s, 使得对于所有的 $z\in \mathrm{P}$ 有

$$s(z) = \frac{az+b}{cz+d},$$

并且用这种方式 (让 a, b, c, d 变化) 得到的 P 的置换组成集合 P 的一个变换群.

c) 证明在 b) 定义的从 P 到 P 内的映射 s 可以分解为属于下列类型之一的变换的乘积:

$$z \to z+u \quad (u \text{ 是实数}); \quad z \to vz \quad (v \text{ 是正实数}); \quad z \to -1/z.$$

给出这些映射的几何解释.

d) 如果附标为 z 的点描绘一个包含于 P 内的圆周, 或者一个中心在实轴上的包含于 P 内的半圆周, 或者一条起点在实轴并且垂直于实轴的半直线, 问 $s(z)$ 描绘的图形是什么?

¶¶ 11. 设 G 是由

$$s(z) = \frac{az+b}{cz+d}$$

给定的 Poincaré 半平面 P 的置换的集合, 其中 a, b, c, d 是满足条件 $ad - bc = 1$ 的有理整数 (见习题 10).

a) 证明 G 是 P 的一个置换群 (称为**算术模群**), 它含有由

$$u(z) = z + 1, \quad v(z) = -1/z$$

给定的映射 u 和 v. 在下面我们打算证明 G 是 u 和 v 生成的. 用 G_0 表示由 u 和 v 生成的 G 的子群.

b) 对于所有的 $z \in P$, 设 I_z 是具有下列性质的实数 $y > 0$ 的集合: 存在一个 $s \in G_0$, 使得 $y = \text{Im}(s(z))$. 利用习题 10 的问题 a), 证明对于所有 $m > 0$, I_z 的满足条件 $y > m$ 的元素的数目是有限的.

c) 证明对于所有 $z \in P$, 存在 $s \in G_0$, 使得数 $z_0 = s(z)$ 满足条件

$$\text{Im}(t(z_0)) \leqslant \text{Im}(z_0) \quad 对于所有\ t \in G_0.$$

证明

$$|z_0| \geqslant 1, \quad -\frac{1}{2} \leqslant \text{Re}(z_0) \leqslant \frac{1}{2}.$$

d) 设 D 是上面的不等式定义的 P 的子集 (由此得到 P 的所有的元素在 G_0 作用下的轨道和 D 相遇). 设 z 是 D 的一个所谓 "内部的" 点, 即它满足条件

$$|z_0| > 1, \quad -\frac{1}{2} < \text{Re}(z_0) < \frac{1}{2}.$$

证明仅有一个 $s \in G$, 使得 $s(z) \in D$ 是 G 的中性元. 对于所有 $t \in G$, 证明存在一个 $t' \in G_0$, 使得 $t'(t(z)) \in D$. 由此推出 $t \in G_0$, 因此有 $G = G_0$.

e) 证明对于所有 $s \in G$, 集合 $s(D)$ 是一个由垂直于实轴的半圆周界定的所谓的 "三角形" (有一些半圆周可能退化为半直线), 指出半平面 P 是 $s(D) (s \in G)$ 的并集, 并且两个不同的区域 $s(D), t(D)$ (即对于它们 $s \neq t$) 不可能有公共边. 用直尺和圆规以最大可能的精确度做由半平面的这个 "铺砌" 组成的图形 —— 从 D 出发, 然后构造集合 $u^n(D)$, 再后构造集合 $vu^n(D)$, 再后构造集合 $u^p vu^n(D)$, 等等. [如果持续构造充分多次, 那么所得到的图形是极其复杂的, 该图形可以作为 F. Klein 和 H. Poincaré 关于 "自守函数" 的工作的出发点.]

12. 设 K 是 $\mathbf{C}$ 的由形如 $x + iy$ (称为 **Gauss 整数**) 的数组成的子环, 其中的 $x, y \in \mathbf{Z}$.

a) 设 $u, v \in K$, $v \neq 0$; 令

$$u/v = x + iy,$$

其中的 x, y 是有理数, 再确定有理整数 m 和 n, 使得

$$|x - m| \leqslant \frac{1}{2}, \quad |y - n| \leqslant \frac{1}{2};$$

最后令

$$q = m + in, \quad r = u - qv.$$

证明有 $\mathrm{N}(r) < \mathrm{N}(v)$, 即 $|r| < |v|$.

b) 由此推出环 K 的所有理想都是主理想 (仿照 §7 的例 8 中所使用的推理).

13. 对于所有的素数 p, 用 $\mathbf{F}_p$ 表示域 $\mathbf{Z}/p\mathbf{Z}$. 求 $d \in \mathbf{F}_p$, 使得环 $\mathbf{F}_p[\sqrt{d}]$ 对于 $p = 2, 3, 5, 7, 11$ 是一个域. 你发现与 §8 的习题 11 的关系了吗?

14. 设 K 是一个交换域, u 和 v 是 K 的元素, 它们在 K 内不是平方. 存在从域 $\mathrm{K}[\sqrt{u}]$ 到域 $\mathrm{K}[\sqrt{v}]$ 上的一个同构, 使得对于所有 $x \in \mathrm{K}$ 有 $f(x) = x$, 必须并且只需 u/v 是 K 的一个元素的平方.

应用: 考虑习题 13 中构造的域, 指出任意两个这样的域是同构的, 只要它们的元素数目相同. (这个结果是下列事实的一个特殊情形: 有同样多的元素的有限域总是同构的.)

15. 设 K 是一个交换域, 而 d 是 K 的一个元素, 它在 K 内不是一个平方. 假定 K 的元素 $2 = 1 + 1$ 不是零 (这就排除了比如域 $\mathbf{Z}/2\mathbf{Z}$). 指出, 域 $\mathrm{K}[\sqrt{d}]$ 的不属于 K 的元素 x 在 $\mathrm{K}[\sqrt{d}]$ 内是一个平方, 必须并且只需 $\mathrm{N}(x)$ 在 K 内是一个平方.

应用: 取 $\mathrm{K} = \mathbf{Z}/11\mathbf{Z}$ 和 $d = 7$. 求在 $\mathrm{K}[\sqrt{d}]$ 内是平方的所有元素 $u + v\sqrt{d}(u, v \in \mathrm{K})$ (可以把当 u 和 v 取遍 K 时 $u^2 - 7v^2$ 的值集中在一个行列对查表内). 由此得到一个 14641 个元素的有限域的例子, 并且验证你所得到的所有的域是两两同构的.

16. 设 K 是一个交换域, 其中的 -1 是一个平方. 设 d 是 K 的一个元素, 它不是 K 内的一个平方. 指出 $\sqrt{d}$ 在 $\mathrm{K}[\sqrt{d}]$ 内不是一个平方.

应用: 构造一个有 17^4 个元素的域.

¶¶ 17. 设 K 是有 q 个元素的有限域, 假定 q 是奇数.

a) 证明 K 的元素 1 和 -1 是不同的 (注意到在相反情形将有对于所有的 $x \in \mathrm{K}$ 有 $2x = qx = 0$, 并且注意到 2 和 q 是互素的).

b) 证明 $x \to x^2$ 是从 K 的非零元素的乘法群 K^* 到自身内的一个同态, 其核由 1 和 -1 组成. 由此推出 K 中非零平方组成一个有 $\dfrac{q-1}{2}$ 个元素的 K^* 子群 H (利用 §7 的习题 21), 并且商群 K^*/H 有两个元素.

c) 证明对于所有的 $x \in \mathrm{K}^*$ 有

$$x^{\frac{q-1}{2}} = \begin{cases} 1, & \text{如果 } x \text{ 是一个平方}, \\ -1, & \text{如果 } x \text{ 不是一个平方}. \end{cases}$$

d) 设 p 是一个奇素数. 我们说一个不被 p 整除的整数 n 是一个模 p 的**二次剩余**, 如果存在一个 $x \in \mathbf{Z}$, 使得

$$x^2 \equiv n (\bmod\ p).$$

证明二次剩余分成模 p 的 $\dfrac{p-1}{2}$ 个类, 并且对于所有不被 p 整除的整数 n 有

$$n^{\frac{p-1}{2}} = \begin{cases} 1, & \text{如果 } n \text{ 对于模 } p \text{ 是二次剩余}, \\ -1, & \text{如果 } n \text{ 对于模 } p \text{ 不是二次剩余} \end{cases}$$

(Euler 判别法). 证明 -1 是模 p 二次剩余当且仅当

$$p \equiv 1 (\bmod\ 4).$$

[在 18 世纪和 19 世纪初二次剩余理论的研究是整个 “近代” 数论和代数的出发点之一. 这条道路上的最著名的结果是 Euler 猜想的**二次互反律**, 并且被 Gauss 在 19 岁时证明了, 堪称一个美丽的开端. 为了叙述它, 应当引进 **Legendre 符号**

$$\left(\frac{n}{p}\right)=\begin{cases}1, & \text{如果 } n \text{ 是模 } p \text{ 的二次剩余,}\\ -1, & \text{如果 } n \text{ 不是模 } p \text{ 的二次剩余.}\end{cases}$$

那么二次互反律是, 如果 p 和 q 是不同的奇素数, 则有

$$\left(\frac{p}{q}\right)\cdot\left(\frac{q}{p}\right)=(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}.$$

对于这个结果以及一般的初等数论, 参见例如 G. H. Hardy 和 E. M. Wright 的 *An Introduction to the Theory of Numbers* (Oxford, 1938).

在 Gauss 之后, 数学家致力于推广二次互反律到代数整数环 (§34, 习题 48), 这比 Gauss 所处理的有理整数的情形无可比拟的困难. 完整的结果组成 “类域论”, 在被 Hilbert 猜想之后, 由 Takagi, Artin 和 Hasse 在 1920—1925 年获得. 十几年之后这个理论又被 Chevalley 大大地改进, 并且目前仍然是十分活跃的研究对象.]

第三章　环上的模

一个环上的模的概念, 提供了处理 "线性" 问题的一个纯代数框架, 在数学的所有分支里都会遇到这类问题: 数论, 古典线性代数, 张量计算, 微分形式, 偏微分方程, 积分方程, 代数几何, 解析函数, 代数拓扑, 等等.

一个环 K 上的模的理论的某些结果仅在关于环 K 的某些假设成立时才是有效的. 例如, "维数" 理论需要假定 K 是一个域.

反之, 最简单的结果对于*所有的*环都成立. 在本章 (§10 至 §17) 看到的就是这些结果. 读者或许发现这里采用的观点太一般太抽象, 他可能更喜欢基础环 K 就是, 比方说, 实数域 $\mathbf{R}$, 但是这样做并不能够达到实质性的简化所提到的各节的表述: 如果假设 $K = \mathbf{R}$, 可以忽略 §10 的例 5, 6, 9, 10 和第 4 小节, §11 的例 4, 5, 6, 11, §16 的注 4, 以及 §17 的例 2, 所有其余的内容除了把字母 K 换成 $\mathbf{R}$ 以外, 没有任何改变.

§10　模和向量空间

1. 环上的模的定义

设 K 是一个环. 我们称由一个集合 M 和 M 上的一个记为

$$(x, y) \to x + y$$

的运算以及从集合 $K \times M$ 到 M 内的一个记为

$$(\lambda, x) \to \lambda x$$

的运算组成的对象为**环 K 上的左模**或**左 K-模**, 如果 M 和两个运算满足以下两个条件:

(M1) 集合 M 配备了运算 $(x,y) \to x+y$ 是一个交换群;

(M2) 对于任意 $x,y \in \mathrm{M}$ 和 $\lambda,\mu \in \mathrm{K}$ 有关系

$$\lambda(\mu x) = (\lambda\mu)x; \quad 1x = x;$$
$$(\lambda+\mu)x = \lambda x + \mu x; \quad \lambda(x+y) = \lambda x + \lambda y.$$

在模理论中, 环 K 是一次性取定的, 一般称为**基础环**, 其元素取名是**标量** (经常用希腊字母表示); 而 K-模的元素则取名为**向量** (经常用拉丁字母表示, 许多人认为应当在字母上加一个箭头, 众多数学家很久以来已经不这样做了). 但是 "标量" 和 "向量" 的区分并没有任何精确的数学意义 (下面的例 1 事实上指出 "标量" 是特殊的 "向量"), 这样取名的目的无非是帮助读者想象熟悉的几何图形.

当环 K 是一个域时, 我们用 K 上的*左向量空间*代替左 K-模. 特殊情形是, 实数域 $\mathbf{R}$ 上的向量空间称为**实向量空间**, 复数域 $\mathbf{C}$ 上的向量空间称为**复向量空间**. 这两个概念在分析中和在物理学中远非最重要的; 反之, 任意域上的向量空间, 环 $\mathbf{Z}$ 上的 (见下面的例 5) 或 "多项式环" 上的模在数学的许多分支里比实或复向量空间起着更重要的作用. 甚至在理论物理学里, 人们利用环上的模, 这些环不是域也不是交换的 (Lorentz 群的线性表示, 旋量, 等等), 尽管物理学家还没有使用模论的语言.

我们注意公理 (M2) 蕴含关系

$$\lambda 0 = 0 \quad \text{对于任意 } \lambda \in \mathrm{K}; \quad 0x = 0 \quad \text{对于任意 } x \in \mathrm{M}.$$

(在这些关系里, 符号 0 既表示 K 的零元, 也表示加法群 M 的零元. 读者会容易地选择使得写出来的关系有意义的解释.) 为了确立第一个关系, 我们注意对于所有 $x \in \mathrm{M}$ 有 $\lambda 0 + \lambda x = \lambda(0+x) = \lambda x$, 于是得到所宣布的 $\lambda 0 = \lambda x - \lambda x = 0$. 第二个关系基于事实 $0x + 1x = (0+1)x = 1x$, 由此即得 $0x = x - x = 0$.

当然我们使用刚得到的等式以及公理 (M2) 中的关系时, 并不注明引自哪里.

最后注意到右 K-模如下定义: 我们称由一个加法群 M 和从集合 $\mathrm{K} \times \mathrm{M}$ 到 M 内的一个记为

$$(x,\lambda) \to x\lambda$$

的运算组成的对象为**右 K-模**, 如果这个运算满足以下等式所表示的条件:

$$(x\lambda)\mu = x(\lambda\mu); \quad x1 = x;$$
$$x(\lambda+\mu) = x\lambda + x\mu; \quad (x+y)\lambda = x\lambda + y\lambda.$$

容易指出 (第 4 小节) *右* K-模只不过是一个由 K 经过一个简单的步骤演绎出的环上的*左* K-模 (并且如果 K 是交换的, 它等同于 K, 以致两个概念的区别仅对于非交换环才有意义). 这样我们既使用左模的语言, 也使用右模的语言, 不言而喻, 经过平凡的翻译就从一个过渡到另一个.

现在举几个模和向量空间的重要例子.

2. 模的例子

例 1 对于所有的环 K 和整数 $n \geqslant 1$, 考虑集合

$$\mathrm{K}^n = \mathrm{K} \times \cdots \times \mathrm{K} \ (n \text{ 个因子}),$$

而作为左 K-模, 取

$$(\xi_1, \cdots, \xi_n) + (\eta_1, \cdots, \eta_n) = (\xi_1 + \eta_1, \cdots, \xi_n + \eta_n),$$
$$\lambda \cdot (\xi_1, \cdots, \xi_n) = (\lambda\xi_1, \cdots, \lambda\xi_n);$$

(M1) 满足这一事实从 §7 第 2 小节推出 (群的直积), 利用环的公理容易验证模的公理 (M2) 中出现的等式.

对于 $n = 1$, 在前面构造中可以把 K 本身看作一个左 K-模 (这表明 "标量" 也是 "向量").

自然也可以把 K^n 看作右 K-模: 加法如前定义, 标量乘法如下定义:

$$(\xi_1, \cdots, \xi_n) \cdot \lambda = (\xi_1\lambda, \cdots, \xi_n\lambda).$$

这是一个*右* K-模 K^n, 它自然地出现在我们将会考察的线性方程组理论中 (但是再说一次, 如果 K 是交换的, 区别失去意义).

例 2 取 $\mathrm{K} = \mathbf{R}$ 为实数域, 而 M 是通常空间的给定起点为 O 的所有向量的集合. 由平行四边形规则定义两个向量的和, 而一个起点为 O 向量 x 和一个实数 λ 的乘积由 x 经受中心为 O 比例为 λ 的位似变换得到, 这样我们得到一个实向量空间 M.

当然, 这里应当证明公理 (M1) 和 (M2) 满足, 这只要在初等几何的框架内就能实现 (理由是我们这里没有给出通常的 "向量" 概念的任何严格的数学定义).

这个例子显然是促使向量空间和模的概念诞生的例子之一, 并且解释了使用向量这个词称呼模的元素的缘由.

例 3 在带两个坐标轴 Ox 和 Oy 的平面上, 考虑起点为 O 在所选定的坐标系里分量为有理数的向量的集合 M. 显然如果 $x, y \in \mathrm{M}$, 则有 $x + y \in \mathrm{M}$, 并且如果 $x \in \mathrm{M}$, $\lambda \in \mathbf{Q}$, 则有 $\lambda x \in \mathrm{M}$. 这样就可以把 M 看作有理数域 $\mathbf{Q}$ 上的向量空间.

例 4 设 K 是一个环, 而 M 是一个左 K-模 (比如 K 自己), 而 X 是任意一个集合. 用 E 表示所有映射

$$f : \mathrm{X} \to \mathrm{M}$$

的集合. 我们要把 E 做成一个左 K-模. 为此应当定义从 X 到 M 内的两个映射的和 $f + g$, 这将是函数 $f(x) + g(x)$, 它在每个 $x \in \mathrm{X}$ 的值由 f 和 g 在 x 的值 (在 M 内)

相加得到. 还应当定义一个数 $\lambda \in \mathrm{K}$ 和一个从 X 到内 M 的映射 f 的乘积 λf: 这将是函数 $\lambda f(x)$, 它在每个 $x \in \mathrm{X}$ 的值由 f 在 x 的值乘以 λ 得到.

留给读者作为习题仔细验证出现在模的定义里的条件 (M1) 和 (M2).

我们注意如果 $\mathrm{M} = \mathrm{K}$, 而 $\mathrm{X} = \{1, 2, \cdots, n\}$, 从 X 到 M 内的一个映射正是 K 的元素的一个序列 $(\xi_1, \cdots, \xi_n)$, 即 $\xi_1 = f(1), \cdots, \xi_n = f(n)$, 我们又回到例 1 中的模 K^n.

例 5 我们指出所有的交换群 G 可以看作左 **Z**-模. 为此, 对 G 采用加法记号, 这已经允许在 G 内定义两个元素的加法, 公理 (M1) 就平凡地被验证. 剩下的是定义一个 $n \in \mathbf{Z}$ 和一个 $x \in \mathrm{G}$ 的乘积, 这可以像在 §7 中那样做, 即

$$nx = \begin{cases} x + \cdots + x(n \text{ 项}), & n \geqslant 1, \\ 0, & n = 0, \\ (-n)(-x), & n \leqslant -1. \end{cases}$$

公理 (M2) 归结为 §7 的例 9 和注 1 中建立的运算规则.

如果群 G 采用乘法记号, 当然就必须定义 G 内两个元素 x 和 y 的和为 xy, 而一个 $x \in \mathrm{G}$ 和一个有理整数 n 的乘积为 x^n. 这跟前面没有任何差别, 如果不计较所采用的记号.

最后, 容易验证**所有 Z**-模采用上述方法从一个加法群出发得到.

这个例子表明模理论当中包括了交换群理论, 而向量空间理论 (更不用说实向量空间理论) 则不然.

例 6 设 K 是环 L 的一个子环, 那么可以把 L 看作一个左 K-模, 只要借助 L 上给定的加法和乘法定义模的运算. 换句话说, 如果考虑 L 的两个元素 x 和 y, 那么它们作为 "向量" 的和就定义为它们作为 L 的元素的和, 而 "标量" λ 和 "向量" x 的乘积将是在环 L 里 λ 乘以 x 的乘积. 公理 (M1) 是满足的, 因为如果不考虑 L 里的乘法运算, 它将变为一个加法群; 至于公理 (M2), 它显然从在环 L 内的交换性和分配性规则推断出来.

例如, 可以把域 **R** 看作 **Q** 上的向量空间, 而域 **C** 看作是 **R** 上的或 **Q** 上的向量空间.

例 7 取 $\mathrm{K} = \mathbf{R}$, 并且组成所有处处连续的映射

$$f : \mathbf{R} \to \mathbf{R}$$

(一个实变量的实值函数) 的集合 M. 在分析里证明了如果 f 和 g 是两个这样的函数, 则 $f + g$ 也是处处连续的, 从而属于 M. 并且如果 $f \in \mathrm{M}$, 则对于所有的数 $\lambda \in \mathbf{R}$ 有 $\lambda f \in \mathrm{M}$ (函数 $f + g$ 和 λf 像在上面例 4 中那样定义). 可以直接看出由集合 M,

从 $\mathrm{M} \times \mathrm{M}$ 到 M 内的映射 $(f, g) \to f + g$ 以及从 $\mathbf{R} \times \mathrm{M}$ 到 M 内的映射 $(\lambda, f) \to \lambda f$ 组成的三元组是一个实向量空间.

这个例子 (它是在分析中引进向量空间的起源) 可以经过众多的变化, 代替要求所考虑的函数处处连续, 可以要求它们是在一个给定的点连续的, 或在一个给定的点可导的, 或处处有连续二阶导数的, 等等.

3. 子模, 向量子空间

设 M 是环 K 上的一个左模. 我们称 M 的所有子集 M' 为 M 的**子模**, 如果它满足以下两个条件:

(i) M' 是加法群 M 的一个子群;

(ii) 关系 $x \in \mathrm{M}'$ 和 $\lambda \in \mathrm{K}$ 蕴含 $\lambda x \in \mathrm{M}'$.

为了验证 M 的一个子集 M' 为一个子模, 应当验证 M' 是非空的 (事实上我们验证 $0 \in \mathrm{M}'$) 并且有

$$\lambda x + \mu y \in \mathrm{M}', \quad 任意\ \lambda, \mu \in \mathrm{K} \quad 和\ x, y \in \mathrm{M}'.$$

这个条件显然是必要的. 反之, 假定它满足, 令 $\mu = 0$ 就已经得上面的到条件 (ii). 为了得到条件 (i) , 只需令 $\lambda = 1, \mu = -1$, 并且注意到在一个模里有

$$-x = (-1)x \quad 对于所有\ x \in \mathrm{M}$$

(事实上, $(-1)x + x = (-1)x + (1)x = (-1+1)x = 0x = 0$).

一个模 M 至少总有两个子模, 即 M 和缩减为仅有的向量 0 的集合.

设 M' 为模 M 的一个子模, 上面的条件 (i) 已经允许把 M' 看作一个加法群; 此外条件 (ii) 允许定义从 $\mathrm{K} \times \mathrm{M}'$ 到 M' 内的一个映射 $(\lambda, x) \to \lambda x$, 而出现在模的公理 (M2) 里的等式在 M 内是成立的, 在 M' 里更不必说是成立的. 因此可以把 M 的所有子模看作一个左 K-模.

当 K 是一个域时, 我们说子向量空间以代替子模.

例 8　取 M 为例 2 中的实向量空间, 那么在 M 内 (在 $\{0\}$ 和 M 本身之外) 有两类子向量空间: a) 起点为 O 位于过 O 的一条直线上的向量的集合; b) 起点为 O 位于过 O 的一张平面上的向量的集合. 此外容易看出除了这两类, M 再没有其他的子向量空间.

例 9　设 K 是一个环, 并且把它看作 (例 1) 一个左 K-模, 那么它的子模是 K 的子集 I, 它们是非空的, 并且

$$ux + vy \in \mathrm{I}, \quad 任意\ u, v \in \mathrm{K}\ 和\ x, y \in \mathrm{I},$$

于是这些是在 §8 第 6 小节定义的 K 的左理想.

例 10 设 G 是一个采用加法记号的交换群, 并且把 G 看作一个 **Z**-模 (例 5), 那么 G 的子模正是它的子群, 因为如果 H 是 G 的一个子群, 那么对于所有 $x \in \mathrm{H}$ 和所有 $n \in \mathbf{Z}$ 有 $nx \in \mathrm{H}$.

以下结果会经常用到:

定理 1 *设 L 是一个环上的模, 而 $(\mathrm{M}_i)_{i\in \mathrm{I}}$ 是 L 的子模的一个族. 那么诸 M_i 的交集是还是 L 的一个子模. 而为了诸 M_i 的并集是 L 的一个子模, 只需对于任意的 $i, j \in \mathrm{I}$ 存在一个 $k \in \mathrm{I}$ 使得 $\mathrm{M}_i \subset \mathrm{M}_k$, 并且 $\mathrm{M}_j \subset \mathrm{M}_k$.*

这个结果严格按照 §7 的定理 1 来证明, 我们让读者撰写严格详细的证明.

不言而喻, 子模的并集一般不是子模: 例如, 在古典情形 (例 8), 两条过原点的不同直线的并集不是一张平面.

4. 右模和左模(*)

设 K 是一个环. 我们要构造一个新的环, 称为 K 的**反环**, 记作

$$\mathrm{K}^{\mathrm{o}};$$

像我们就要指出的那样, K 上的*右*模正是 K^{o} 上的*左*模.

为了构造 K^{o}, 应当给定一个集合和其上的两个运算, 即一个 "加法" 和一个 "乘法". 作为定义, 集合 K^{o} 就是集合 K (环 K 和环 K^{o} 有相同的元素), 而 K^{o} 上的*加法*就是 K 上的*加法* (K 的两个元素和 $x+y$ 的值不论在 K 里计算还是在 K^{o} 里计算有相同的值). 另一方面, 在 K^{o} 里的乘法不再是 K 里给定的乘法 $(x, y) \to xy$, 而是映射 $(x, y) \to yx$. 换句话说, 如果用 xy 表示环 K 内两个元素的乘积, 而用 $x * y$ 表示在环 K^{o} 里它们的乘积, 则有关系

$$x * y = yx.$$

容易看出集合 $\mathrm{K}^{\mathrm{o}}(=\mathrm{K}!)$ 配备了刚定义的两个运算是一个环. 例如, 公式

$$(x+y) * z = x * z + y * z$$

显然归结为在环 K 里的 $z(x+y) = zx + zy$.

不言而喻, 如果 K 是一个交换环, 环 K^{o} 等同于环 K. 仅在非交换的情形 K^{o} 的构造才有兴趣.

现在设 M 是环 K 上的一个*右*模. 令

$$\lambda * x = x\lambda \quad \text{对于所有 } x \in \mathrm{M} \text{ 和所有 } \lambda \in \mathrm{K},$$

(*) 这一小节初读可以略去, 仅在 §16 的末尾用到它, 而初学的读者现在可以假定环 K 是交换的.

就定义了从 $K^o \times M$ 到 M 内的一个映射. 那么加法群 M 配备了刚才定义的映射就是 K 的反环 K^o 上的一个左模. 事实上, 我们有

$$\lambda * (x+y) = (x+y)\lambda = x\lambda + y\lambda = \lambda * x + \lambda * y,$$
$$(\lambda+\mu) * x = x(\lambda+\mu) = x\lambda + x\mu = \lambda * x + \mu * x,$$
$$\lambda * (\mu * x) = (\mu * x)\lambda = (x\mu)\lambda = x(\mu\lambda) = (\mu\lambda) * x = (\lambda * \mu) * x,$$

最后还有

$$1 * x = x1 = x,$$

这就证明了所陈述的结果.

我们刚陈述的构造表明对于左模 (对应的, 右模) 证明的结果自动应用到右模 (左模), 只需从给定的基础环过渡到它的反环. 还可以发现当基础环是交换环时, 在模理论里, 把标量放在向量左边抑或右边, 是完全没有区别的, 这纯粹是写法约定的问题, 它跟数学事实本身没有任何关系, 而人们受制于不能够从 "右" 写法过渡到 "左" 写法和反向的过渡, 完全是作茧自缚.

(本节习题在 §11 之后.)

§11 模内的线性关系

1. 线性组合

设 $a_1, \cdots, a_n$ 是环 K 上一个左模 M 的元素, 称具有下列性质的所有向量 $x \in M$ 为 $a_1, \cdots, a_n$ 的一个**线性组合**: 存在 $\xi_1, \cdots, \xi_n \in K$ 使得

$$x = \xi_1 a_1 + \cdots + \xi_n a_n.$$

对于右 K-模必然有一个类似的概念.

例 1 在 K^n (§10, 例 1) 里考虑向量

$$e_1 = (1, 0, 0, \cdots, 0, 0),$$
$$e_2 = (0, 1, 0, \cdots, 0, 0),$$
$$\cdots\cdots\cdots\cdots$$
$$e_n = (0, 0, 0, \cdots, 0, 1).$$

显然有

$$\xi_1 e_1 = (\xi_1, 0, 0, \cdots, 0, 0),$$
$$\xi_2 e_2 = (0, \xi_2, 0, \cdots, 0, 0),$$
$$\cdots\cdots\cdots\cdots$$
$$\xi_n e_n = (0, 0, 0, \cdots, 0, \xi_n);$$

把得到的结果相加即得

$$\xi_1 e_1 + \cdots + \xi_n e_n = (\xi_1, \cdots, \xi_n). \tag{1}$$

换句话说, K^n 的**所有**元素都是向量 $e_1, \cdots, e_n$ 的线性组合; 甚至有更精确的结果: 给定向量 $x \in \mathrm{K}^n$, 存在唯一的一组数 $\xi_1, \cdots, \xi_n$ 使得

$$x = \xi_1 e_1 + \cdots + \xi_n e_n.$$

例 2 考虑右 K-模 K^p, 其中 $p \geqslant 1$, 和这个模的给定向量:

$$\begin{aligned} a_1 &= (\alpha_{11}, \alpha_{21}, \cdots, \alpha_{p1}), \\ a_2 &= (\alpha_{12}, \alpha_{22}, \cdots, \alpha_{p2}), \\ &\cdots\cdots\cdots\cdots \\ a_n &= (\alpha_{1n}, \alpha_{2n}, \cdots, \alpha_{pn}). \end{aligned}$$

设 $b = (\beta_1, \cdots, \beta_p)$ 是 K^p 的一个向量, 那么关系

$$b = a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n \tag{2}$$

等价于如下的 n 个未知量 $\xi_1, \cdots, \xi_n, p$ 个线性方程的方程组:

$$\begin{cases} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \cdots + \alpha_{1n}\xi_n = \beta_1, \\ \qquad\cdots\cdots\cdots\cdots \\ \alpha_{p1}\xi_1 + \alpha_{p2}\xi_2 + \cdots + \alpha_{pn}\xi_n = \beta_p. \end{cases} \tag{3}$$

事实上, 我们有

$$\begin{aligned} a_1\xi_1 &= (a_{11}\xi_1, \cdots, a_{p1}\xi_1), \\ &\cdots\cdots\cdots\cdots \\ a_n\xi_n &= (a_{1n}\xi_n, \cdots, a_{pn}\xi_n), \end{aligned}$$

故关系 (3) 的左端就是关系 (2) 右端的分量(*).

这个例子是线性方程组的 "几何" 理论的出发点.

定理 1 设 $a_1, \cdots, a_n$ 是一个左 K-模 M 的元素, 而 M′ 是所有 $a_1, \cdots, a_n$ 的线性组合的集合. 则 M′ 是含有 $a_1, \cdots, a_n$ 的 M 的最小子模.

首先, 关系

$$a_1 = 1 \cdot a_1 + 0 \cdot a_2 + \cdots + 0 \cdot a_n$$

和对于 $a_2, \cdots, a_n$ 的类似关系表明 M′ 含有 $a_1, \cdots, a_n$. 其次, 对于任意 $\xi_1, \cdots, \xi_n \in$ K, M 的所有含有 $a_1, \cdots, a_n$ 的子模含有 $\xi_1 a_1, \cdots, \xi_n a_n$, 因此含有 $\xi_1 a_1 + \cdots + \xi_n a_n$. 于是含有 a_i $(1 \leqslant i \leqslant n)$ 的 M 的所有子模包含 M′.

(*) 在 K^n 里, 称标量 $\xi_1, \cdots, \xi_n$ 为向量 $(\xi_1, \cdots, \xi_n)$ 的分量.

最后, 为了完成证明, 剩下要做的事情是验证 M′ 实际上是 M 的一个子模. 设

$$x = \xi_1 a_1 + \cdots + \xi_n a_n, \quad y = \eta_1 a_1 + \cdots + \eta_n a_n$$

是 M′ 的两个元素, 一个平凡的计算指出

$$\lambda x + \mu y = \zeta_1 a_1 + \cdots + \zeta_n a_n,$$

其中

$$\zeta_1 = \lambda \xi_1 + \mu \eta_1, \quad \cdots, \quad \zeta_n = \lambda \xi_n + \mu \eta_n,$$

由此得到对于任意标量 λ 和 μ 有 $\lambda x + \mu y \in \mathrm{M}'$, 这就完成了证明.

定理 1 中的 M′ 称为 M **的由** $a_1, \cdots, a_n$ **生成的子模**; 当 $\mathrm{K} = \mathbf{Z}$, M 只不过是采用加法记号的交换群, M′ 正好是 M 的由其子集 $\mathrm{B} = \{a_1, \cdots, a_n\}$ 生成的子群 (§7, 第 4 小节).

2. 有限生成模

设 M 是一个左 K-模, 而 M′ 是 M 的一个子模. 我们说 M′ 是**有限生成的**, 如果存在有限个向量 $a_1, \cdots, a_n \in \mathrm{M}'$, 使得 M′ 是由这些向量生成的, 那么我们说 $a_1, \cdots, a_n$ 组成 M′ 的一个**生成元组**.

这个定义特别应用到环 M 本身, 换句话说, 我们说一个模是有限生成的, 如果它含有有限个向量 $a_1, \cdots, a_n$, 使得所有 $x \in \mathrm{M}$ 是 $a_1, \cdots, a_n$ 的线性组合.

当基础环 K 是域时, 则称为**有限维向量空间**, 以代替有限生成模.

例 3 例 1 指出 K^n 是一个有限生成的模.

例 4 当 $\mathrm{K} = \mathbf{Z}$ 时有限生成模的概念归结为在 §7 第 4 小节引进的有限生成群的概念.

例 5 作为 **Z**-模 **Q** 不是有限生成的 (§7, 例 10); 反之, 作为 **Q** 上的向量空间 **Q** 是有限生成的, 见上面的例 3, 其中取 $\mathrm{K} = \mathbf{Q}, n = 1$.

例 6 设 K 是一个环, 我们考虑 K (看作左 K-模) 的有限生成子模, 这是 K 的左理想 I (§10, 例 9), 它具有以下性质: 存在 I 的有限个元素 $a_1, \cdots, a_n$, 使得所有 $x \in \mathrm{I}$ 对于适当选取的 $u_1, \cdots, u_n \in \mathrm{K}$, 可以写成形式

$$x = u_1 a_1 + \cdots + u_n a_n.$$

一个这样的理想称为环 K 的**有限生成左理想**. 由这个例子显然推知 K 的所有主理想是有限生成的, 但其逆一般不真.

给定 K 的元素 $a_1, \cdots, a_n$, 可以写成形式 $u_1 a_1 + \cdots + u_n a_n$ 的元素的集合 (换句话说, K 的由 $a_1, \cdots, a_n$ 生成的子空间) 称为 K **的由** $a_1, \cdots, a_n$ **生成的左理想**.

例 7 在 §10 例 7 所描述的实向量空间不是有限维的.

3. 线性关系

给定一个左 K-模 M 的元素 $a_1, \cdots, a_n$, 设 x 是这些向量的一个线性组合. 考虑把 x 写成给定向量的线性组合的两种方式

$$x = \xi_1 a_1 + \cdots + \xi_n a_n = \eta_1 a_1 + \cdots + \eta_n a_n;$$

通过做差, 立刻得到关系

$$(\xi_1 - \eta_1)a_1 + \cdots + (\xi_n - \eta_n)a_n = 0.$$

这就解释了引进以下定义的理由.

称模 K^n 的所有元素 $(\lambda_1, \cdots, \lambda_n)$ 为向量 $a_1, \cdots, a_n$ 之间的**线性关系**, 如果有

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0,$$

说线性关系 $(0, \cdots, 0)$ 是**平凡的**. 我们说向量 $a_1, \cdots, a_n$ 是**线性无关的**, 或说族 $\{a_i\}_{1 \leqslant i \leqslant n}$ 是**自由的**, 如果除了平凡的线性关系, 不存在 $a_1, \cdots, a_n$ 之间的其他线性关系.

于是说向量 $a_1, \cdots, a_n$ 是线性无关的, 就是说

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0 \quad \text{蕴含} \quad \lambda_1 = \cdots = \lambda_n = 0;$$

反之, 就说它们不是线性无关的 (或说向量 $a_1, \cdots, a_n$ 是线性相关的), 意即存在不全为零的标量 $\lambda_1, \cdots, \lambda_n$, 使得

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0.$$

线性无关的向量 $a_1, \cdots, a_n$ 必然是两两不同的, 因为如果比如说 $a_1 = a_2$, 那么 K^n 的元素 $(1, -1, 0, \cdots, 0)$ 显然是 $a_1, \cdots, a_n$ 之间的一个非平凡的线性关系. 但诸 a_i 两两不等对于它们线性无关是不充分的.

我们引进线性无关概念的方式直接证明了以下结果:

定理 2 给定一个左 K-模 M 的元素 $a_1, \cdots, a_n$, 设 x 是向量 $a_1, \cdots, a_n$ 的一个线性组合. 下列性质是等价的:

a) 仅存在一个 $(\xi_1, \cdots, \xi_n) \in \mathrm{K}^n$, 使得

$$x = \xi_1 a_1 + \cdots + \xi_n a_n;$$

b) $a_1, \cdots, a_n$ 是线性无关的.

只需注意到, 如果 $(\lambda_1, \cdots, \lambda_n)$ 是 $a_1, \cdots, a_n$ 之间的一个线性关系, 则有

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0,$$

因此有

$$\xi_1 a_1 + \cdots + \xi_n a_n = (\xi_1 + \lambda_1)a_1 + \cdots + (\xi_n + \lambda_n)a_n.$$

事实上, 这个性质刻画了 $a_1, \cdots, a_n$ 之间的线性无关性.

例 8 在 K^n 内, 例 1 中的 $e_1, \cdots, e_n$ 是线性无关的.

例 9 取 $\mathrm{K} = \mathbf{R}$ 设 M 为普通空间内的起点为 O 的向量组成的向量空间 (§10, 例 2). 向量 $a_1, \cdots, a_n \in \mathrm{M}$ 是线性无关的, 必须且只需: (1) 当 $n = 1$ 时, a_1 不是零; (2) 当 $n = 2$ 时, 它们不在同一条直线上; (3) 当 $n = 3$ 时, 它们不位于同一个平面上. 当 $n \geqslant 4$ 时, 向量 $a_1, \cdots, a_n$ 绝不能是线性无关的 (因为如果它们是线性无关的, 那么三个向量 a_1, a_2, a_3 就是线性无关的, 其余的向量, 比如 a_4 将是这三个向量的线性组合, 而这显然与线性无关性相矛盾).

例 10 取 $\mathrm{K} = \mathbf{R}$, 取所有映射 $f : \mathbf{R} \to \mathbf{R}$ 组成的向量空间为 M (§10, 例 4, 其中的 $\mathrm{X} = \mathrm{M} = \mathrm{K} = \mathbf{R}$). 设 $f_1, \cdots, f_n$ 是 M 的元素, 即实变量 t 的实值函数. 对于 $\lambda_1, \cdots, \lambda_n \in \mathbf{R}$, 函数 $f = \lambda_1 f_1 + \cdots + \lambda_n f_n$ 是对于所有 $t \in \mathbf{R}$ 由

$$f(t) = \lambda_1 f_1(t) + \cdots + \lambda_n f_n(t)$$

给定的. $f_1, \cdots, f_n$ 之间的一个线性关系是 n 个实数的一个序列 $(\lambda_1, \cdots, \lambda_n) \in \mathbf{R}^n$, 使得

$$\lambda_1 f(t)_1 + \cdots + \lambda_n f_n(t) = 0 \quad \text{对于所有 } t \in \mathbf{R}.$$

作为例子, 考虑 $n+1$ 个函数 $1, t, t^2, \cdots, t^n$, 这些函数之间的一个线性关系是满足

$$c_0 + c_1 t + \cdots + c_n t^n = 0$$

的 $n+1$ 个实数的组 $c_0, c_1, \cdots, c_n$. 在研究一个代数方程的根时, 我们将看到前面的关系蕴含

$$c_0 = \cdots = c_n = 0,$$

由此得到对于任意的 n, 作为实向量空间 M 的元素, $1, t, t^2, \cdots, t^n$ 是线性无关的.

例 11 由于 $\mathbf{Q}$ 是 $\mathbf{C}$ 的一个子域 (§10, 例 6), 可以认为 $\mathbf{C}$ 是 $\mathbf{Q}$ 上的向量空间. 对于一个 $z \in \mathbf{C}$, 考虑 z 的 $n+1$ 个幂 $1, z, \cdots, z^n$, 我们说在 $\mathbf{C}$ 的这 $n+1$ 元素之间存在一个非平凡的 (系数在 $\mathbf{Q}$ 内的) 线性关系, 意即存在不全为零的有理数 $c_0, c_1, \cdots, c_n$, 使得 z 满足方程

$$c_0 + c_1 z + \cdots + c_n z^n = 0.$$

如果至少对于 n 的一个值这件事成立, 则说 z 是一个**代数数**(*). 在相反的情形 (即 z 不满足任何有理系数的非平凡代数方程) 则说 z 是一个**超越数**, 例如 $\pi = 3.14159\cdots$ 就是这种情形 (其证明是很不容易的).

4. 自由模, 基

设 M 是环 K 上的一个左模. 我们说 M 是**有限生成的自由模**, 如果存在 M 的有限个元素 $a_1, \cdots, a_n$, 它们是线性无关的, 并且生成 M. 这时说 $a_1, \cdots, a_n$ 组成 M 的一个基 (于是一个基是线性无关的生成元的组).

设 $a_1, \cdots, a_n$ 是一个左 K-模 M 的元素, 说它们生成 M, 就是说对于所有 $x \in \mathrm{M}$, 关系

$$x = \xi_1 a_1 + \cdots + \xi_n a_n$$

至少对于一个 $(\xi_1, \cdots, \xi_n) \in \mathrm{K}^n$ 是满足的. 而说 $a_1, \cdots, a_n$ 是线性无关的, 意即对于所有的 $x \in \mathrm{M}$, 上面的关系至多对于一个 $(\xi_1, \cdots, \xi_n) \in \mathrm{K}^n$ 是满足的.

因此, $a_1, \cdots, a_n$ 组成 M 的一个基, 必须且只需对于每个 $x \in \mathrm{M}$, 存在唯一的一个 $(\xi_1, \cdots, \xi_n) \in \mathrm{K}^n$ 使得

$$x = \xi_1 a_1 + \cdots + \xi_n a_n.$$

标量 $\xi_1, \cdots, \xi_n$ 称为 x **关于** M **的基** $a_1, \cdots, a_n$ **的坐标或分量**.

前面的叙述说明 x 的坐标是 x 的在 K 内取值的函数, 把它们记作 $f_1, \cdots, f_n$, 这样对于所有 $x \in \mathrm{M}$ 就有

$$x = f_1(x) a_1 + \cdots + f_n(x) a_n.$$

我们说映射

$$f_i : \mathrm{M} \to \mathrm{K} \quad (1 \leqslant i \leqslant n)$$

是模 M **关于基** $a_1, \cdots, a_n$ **的坐标函数**. 通过关系

$$a_j = 0 \cdot a_1 + \cdots + 0 \cdot a_{j-1} + 1 \cdot a_j + 0 \cdot a_{j+1} + \cdots + 0 \cdot a_n$$

可以直接计算向量 a_j 关于基 $a_1, \cdots, a_n$ 的坐标, 我们得到关系

$$f_i(a_j) = \begin{cases} 1, & \text{如果 } i = j, \\ 0, & \text{如果 } i \neq j. \end{cases}$$

(*) 我们提醒, 像在 §5 第 8 小节曾提到的, 这里定义的代数数的概念跟在中学教学中这个名词表示的概念 (其实就是实数概念) 无任何关系.

可以证明代数数组成复数域 **C** 的一个子域. 正是在 19 世纪 (尤其是 Galois 和德国学派的大数学家: Gauss, Kummer, Jacobi, Lejeune-Dirichlet, Dedekind, Kronneker, Hilbert) 对于代数数的研究促使了近世代数的诞生.

另外, 我们有等式

$$f_i(x+y)=f_i(x)+f_i(y),\quad f_i(\lambda x)=\lambda f_i(x);$$

事实上, 关系

$$x=f_1(x)a_1+\cdots+f_n(x)a_n,\quad y=f_1(y)a_1+\cdots+f_n(y)a_n$$

两端分别相加即得

$$x+y=[f_1(x)+f_1(y)]a_1+\cdots+[f_n(x)+f_n(y)]a_n,$$

此式表明 $x+y$ 的坐标是 $f_i(x)+f_i(y)$. 第二个关系按类似方式建立.

换句话说, 为了两个向量求和, 就把它们相应的分量相加, 而为了一个标量乘以一个向量, 就把这个标量乘以它的每个分量.

例 12 左 K-模 K^n 是有限生成的自由模, 并且例 1 指出向量 $e_1,\cdots,e_n$ 组成这个模的一个基, 称它是 K^n 的**典范基**. 给定 K^n 的一个元素

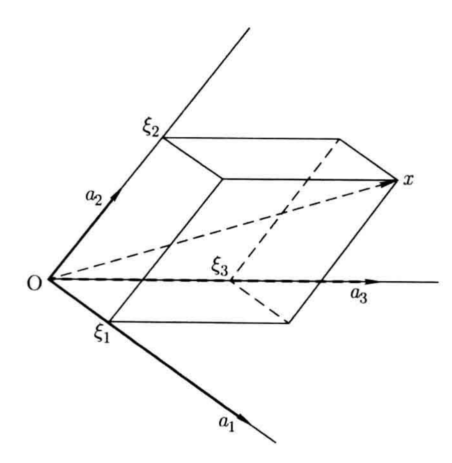
$$x=(\xi_1,\cdots,\xi_n),$$

关系

$$x=\xi_1e_1+\cdots+\xi_ne_n$$

表明 x 的关于 K^n 的典范基的坐标就是数量 $\xi_1,\cdots,\xi_n$.

例 13 再考察上面的例 9. $a_1,\cdots,a_n\in\mathrm{M}$ 组成 M 的一个基, 必须而且只需 $n=3$, 并且向量 a_1,a_2,a_3 不在同一个平面上. 一个向量 x 的坐标按照 "平行六面体" 规则计算, 考察下图即可明白. 利用向量 a_i 来定直线方向, 并且定义这些直线的 "长度单位".



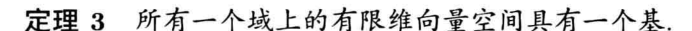
前面的定义尤其可以用到交换群, 只需取 $K = \mathbf{Z}$. 于是可以谈论**有限生成自由交换群**和交换群的**基**. 给定一个交换群 G (采用加法记号, 但是读者把下面叙述的内容翻译成采用乘法记号的形式是有益的), G 的一个基是 G 的元素的一个有限[*]序列 $a_1, \cdots, a_n$, 使得从 $\mathbf{Z}^n$ 到 G 内的映射

$$(r_1, \cdots, r_n) \to r_1 a_1 + \cdots + r_n a_n$$

是双射的, 说 G 是有限生成的自由群, 如果它至少具有一个基.

由于集合 $\mathbf{Z}^n$ 是无限的, 一个有限生成的自由交换群显然必须是无限的. 因此一个有限交换群是一个有限生成的 $\mathbf{Z}$ 模, 但不具有基.

我们发现一个环上的模可以是有限生成的, 却不是有限生成的自由模, 即不具有基. 尽管如此, 我们有:



定理 3 所有一个域上的有限维向量空间具有一个基.

由于本章仅集中注意任意基础环上的结果, 这里我们不证明这个定理. 如果愿意, 鉴于定理 3 的证明仅要求本节的内容, 读者可以直接参考 §19 第 1 小节.

定理 3 表明, 如果 K 是一个任意的环, 断言 "所有有限生成的 K-模是有限生成的自由模" 是错误的, 如果 K 是一个域, 这个断言是正确的.



注 1 为了保证某些陈述的有效性, 我们约定, 对于所有环 K, 缩减为 0 的 K-模是有限生成自由环, 并且具有一个由 0 向量组成的基.

为了明白这个约定, 应当像下面那样定义一个 K-模 M 的基的概念: 这是生成 M 的线性无关向量的一个有限族 $(a_i)_{i\in \mathrm{I}}$; 这个定义允许了指标集 I 是空集, 而要赋予缩减为 0 的模一个基, 这样做是必需的.

同样约定模 K^{o} 是缩减为一个向量 0 的模是必需的.

5. 无穷线性组合[**]

设 K 是一个环, I 是一个任意集合 (有限或无限), 而 $(\lambda_i)_{i\in \mathrm{I}}$ 是以 I 做指标集的 K 的元素的一个族. 如果使得 $\lambda_i \neq 0$ 的 $i \in \mathrm{I}$ 的集合是有限的, 就说标量 λ_i **几乎全是零**, 这个定义以自然的方式推广到任意的模的元素族 $(x_i)_{i\in \mathrm{I}}$.

不言而喻, 刚引进的概念仅当 I 为无限集时才是有兴趣的.

设 $(x_i)_{i\in \mathrm{I}}$ 是 K-模 M 的元素的一个族. 假定 x_i 几乎全是 0, 于是存在 I 的有限子集 J, 使得对于 $i \in \mathrm{I} - \mathrm{J}$ 有 $\lambda_i = 0$. 我们定义

$$\sum_{i\in \mathrm{I}} x_i = \sum_{i\in \mathrm{J}} x_i,$$

显然右端的值不依赖 J 的选择 (只要 J 满足上面的条件).

[*] 但是要参照下一小节.

[**] 本小节在 §26 之前用不到; 读者可以等到需要时再学习.

设 $(x_i)_{i\in\mathrm{I}}$ 是一个 K-模 M 的任意的一个元素族. 称所有具有下述性质的 $x \in \mathrm{M}$ 为 $(x_i)_{i\in\mathrm{I}}$ 的一个**线性组合**: 存在几乎全是零的标量的一个族 $(\xi_i)_{i\in\mathrm{I}}$ 使得

$$x = \sum_{i\in\mathrm{I}} \xi_i x_i,$$

关系有意义, 因为向量 $\xi_i x_i$ $(i \in \mathrm{I})$ 几乎全是零.

容易证明, $(x_i)_{i\in\mathrm{I}}$ 的线性组合的集合 M′ 是含有所有 $x_i, i \in \mathrm{I}$ 的 M 的最小子模, 称为**由族** $(x_i)_{i\in\mathrm{I}}$ **生成的** M **的子模**.

此外, 称所有几乎全是零的标量族 $(\lambda_i)_{i\in\mathrm{I}}$ 为 x_i $(i \in \mathrm{I})$ 之间的一个**线性关系**, 如果

$$\sum_{i\in\mathrm{I}} \lambda_i x_i = 0.$$

如果这个关系蕴含对于所有 $i \in \mathrm{I}$ 有 $\lambda_i = 0$, 则说 x_i $(i \in \mathrm{I})$ 是**线性无关的**, 或说 $x_i (i \in \mathrm{I})$ 是 M 的元素的一个**自由族**.

称生成 M 的 M 的元素的一个自由族 $(x_i)_{i\in\mathrm{I}}$ 为 M 的一个**基**.

如果一个模有一个自由基, 则说它是一个**自由模**.

可以证明域上的所有向量空间具有一个基. 但是这个结果的证明明显地比定理 3 的证明更困难, 并且超越了本书的范围.

例 14 例 10 表明函数

$$t^n \quad (n = 0, 1, 2, \cdots)$$

的无限族在例 10 所提到的向量空间中是自由的. 这些函数的线性组合就是一个实变量的**多项式函数** (将在 §28 研究).

例 15 把 **C** 看作 **Q** 上的一个向量空间, 说一个数 $z \in \mathbf{C}$ 是超越数, 意即 z 的幂的 (无限) 族是自由的.

§§10, 11 习题

1. 证明在 $\mathbf{R}^3$ 内, 向量 $x = (6, 2, -7)$ 是向量

$$a = (2, 1, -3), \quad b = (3, 2, -5), \quad c = (1, -1, 1)$$

的线性组合, 向量 a, b, c 组成 $\mathbf{R}^3$ 的一个基吗?

对于 $\mathbf{R}^4$ 内的向量 $x = (7, 14, -1, 2)$ 和

$$a = (1, 2, -1, -2), \quad b = (2, 3, 0, -1), \quad c = (1, 2, 1, 3), \quad d = (1, 3, -1, 0)$$

解答同样的问题.

在下列的每一个情形, 证明向量 a, b, c 组成 $\mathbf{R}^3$ 的一个基, 并且求向量 x 关于这个基的坐标:

$$a = (1,1,1), \quad b = (1,1,2), \quad c = (1,2,3), \quad x = (6,9,14),$$
$$a = (2,1,-3), \quad b = (3,2,-5), \quad c = (1,-1,1), \quad x = (6,2,-7).$$

对于 $\mathbf{R}^4$ 内的向量

$$a = (1,2,-1,-2), \quad b = (2,3,0,-1), \quad c = (1,2,1,4), \quad d = (1,3,-1,0),$$

和向量 $x = (7,14,-1,2)$ 在 $\mathbf{R}^4$ 内解答同样的问题.

2. 证明 $\mathbf{R}^2$ 的两个向量 (a,b) 和 (c,d) 组成 $\mathbf{R}^2$ 的一个基, 必须并且只需

$$ad - bc \neq 0.$$

3. 设 M 是通常空间内给定起点 O 的向量组成的实向量空间. 设 $\mathrm{M}' \subset \mathrm{M}$ 是起点为 O 的终点位于空间内的一个给定平面的向量的集合. M' 是 M 的一个向量子空间吗?

4. 设 K 是一个环. 考虑在 K^n 的满足关系

$$x_1 + \cdots + x_n = 0$$

的向量 $(x_1, \cdots, x_n)$ 的集合, 这是 K^n 的一个子模吗? 用

$$x_1 + \cdots + x_n = 1$$

代替前面的关系, 解同样的问题.

5. 考虑 $\mathbf{C}^n$ 内的向量

$$x_k = (\xi_{k1}, \cdots, \xi_{kn}) \quad (1 \leqslant k \leqslant r),$$

其坐标满足不等式

$$|\xi_{kk}| > \sum_{\substack{1 \leqslant j \leqslant r \\ j \neq k}} |\xi_{kj}| \quad (1 \leqslant k \leqslant r);$$

证明这些向量在 $\mathbf{C}$ 上是线性无关的.

6. 由向量

$$a = (1,-1,1,0), \quad b = (1,1,0,1), \quad c = (2,0,1,1)$$

生成的 $\mathbf{R}^4$ 的向量子空间是什么?

7. 把有限生成模, 有限生成自由模和这种模的基等概念, 翻译成以乘法描述的交换群的语言 (按照 §10 的例 5, 一个这样的群看作一个 $\mathbf{Z}$-模).

¶8. 设 V 是有理数域 $\mathbf{Q}$ 上的有限维向量空间. 说 V 的一个子集 M 是 V 内的一个**网**, 如果 M 是加法群 V 的一个有限生成子群, 并且 M 包含 V 的一个生成元组.

a) 设 M 是 V 的一个有限生成子群. M 是 V 的一个网, 必须并且只需, 对于所有 $x \in \mathrm{V}$, 存在一个有理整数 $r \neq 0$, 使得 $rx \in \mathrm{M}$. 证明使得 $rx \in \mathrm{M}$ 的 $r \in \mathbf{Z}$ 的集合是环 $\mathbf{Z}$ 的一个 (非零) 理想.

b) 设 p 是一个素数, 而 M 是 V 的一个网. 用 M_p 表示满足下列条件的 $x \in \mathrm{V}$ 的集合: 存在一个不是 p 的倍数的整数 r, 使得 $rx \in \mathrm{M}$. 证明: 如果把 V 看作 §8, 习题 5 的环 $\mathbf{Z}_p$ 上的模, 则 M_p 是 V 的一个子模.

c) 证明, 如果 M 是 V 的一个网, 则有

$$\mathrm{M} = \bigcap_{p \text{ 是素数}} M_p.$$

d) 设 M 和 N 是 M 的两个网. 证明使得

$$\mathrm{M}_p \neq \mathrm{N}_p$$

的素数 p 只有有限个.

(将在 §18 的习题 1 发现网的补充性质.)

9. 设 A 是交换域 K 的一个子环. 假定 K 的所有元素可以表示成形式 uv^{-1}, 其中的 $u, v \in \mathrm{K}, v \neq 0$, 并且 K 是有限生成的 A-模. 证明 $\mathrm{A} = \mathrm{K}$.

¶10. 设 K 是一个环, M 是一个左 K-模, 而 M′ 是 M 的一个子模.

a) 证明

$$x - y \in \mathrm{M}'$$

是集合 M 上的一个等价关系 [称为**模 M′ 同余**, 并且写作

$$x \equiv y (\mathrm{mod}\ \mathrm{M}'),$$

参见 §7, 第 6 小节] .

b) 用 M/M′ 表示 M 关于这个等价关系的商集, 而 p 是从 M 到 M/M′ 上的典范映射. 证明在 M/M′ 上存在唯一的一个左 K-模结构, 使得对于任意 $x, y \in \mathrm{M}$ 和 $\lambda \in \mathrm{K}$ 有

$$p(x+y) = p(x) + p(y), \quad p(\lambda x) = \lambda p(x)$$

(利用 §4 第 3 小节. 对于类似的结构还可以参见 §7 习题 16 和 §8 习题 7). 我们说配备了这个模结构的 M/M′ 是 M 关于子模 M′ 的**商模**.

c) 令 M/M′ 的每一个子模对应于它在映射 p 下的逆像. 证明这样就得到从 M/M′ 的子模到包含 M′ 的 M 的子模集合上的一个双射.

d) 证明如果 M 是有限生成的, 则对于任意 M′, M/M′ 也如是. 如果 M 是有限生成自由的, M/M′ 也是吗?

¶11. 设 M 是环 K 上的一个左模.

a) 对于所有 $x \in \mathrm{M}$, 称使得 $\lambda x = 0$ 的 $\lambda \in \mathrm{K}$ 的集合为 x 的**零化子**. 证明这是 K 的一个左理想.

b) 假定 K 是整环. 证明其零化子不缩减为 0 的 $x \in \mathrm{M}$ 的集合组成 M 的一个子模 T (称 T 是 M 的**扭子模**, 如果 $\mathrm{T} = \{0\}$, 则说 M 是**无扭的**).

c) 证明商模 M/T (习题 10) 是无扭的.

d) 当 $\mathrm{K} = \mathbf{Z}$ 和 $\mathrm{M} = \mathbf{Z}^2/\mathrm{L}$ 时计算 T, 这里 L 是由向量 $(4, 6)$ 生成的 $\mathbf{Z}^2$ 的子群.

¶12. 称环 K 上的一个左模 M 是**扭模**, 如果对于所有 $x \in \mathrm{M}$ 存在一个**非零**标量 $\lambda \in \mathrm{K}$, 使得 $\lambda x = 0$.

a) 设 M 是一个左 K-模, 而 M′ 是 M 的一个子模. 假定 M′ 和 M/M′ 是扭模, 如果 K 是整环, 则 M 也是扭模.

b) 假定 $\mathrm{K} = \mathbf{Z}$. 证明, 一个有限生成的 K-模是扭模, 必须并且只需它是有限的.

13. 设 M 是环 K 上的一个左模. M 的一个子集 (有限或无限) B 称为 M 的**生成元集**, 如果包含 B 的 M 的仅有的子模是整个 M, 或同样的, 如果 M 的每一个元素是 B 的有限个元素的线性组合.

假定 M 是有限生成的. 证明这时可以从 B 抽取 M 的生成元的有限集 (在 M 内选取生成元 x_i 的一个组, 并且借助 B 的元素表示每一个 x_i).

¶¶14. 设 K 是一个交换域, 而 A 是 K 的一个子环. 假定 K 是 A 的**分式域**, 即对于所有的 $x \in \mathrm{K}$, 存在 A 的元素 u 和 v, 使得 $v \neq 0$, 并且

$$x = uv^{-1}.$$

在下面把 K 看作一个 A-模, 说 K 的一个子集 I 是环 A 的一个**分式理想**, 如果它不缩减为 0, 它是 K 的一个子模, 并且存在 K 的一个元素 $d \neq 0$, 使得

$$d\mathrm{I} \subset \mathrm{A}$$

(这里 $d\mathrm{I}$ 表示积 dx 的集合, 其中的 $x \in \mathrm{I}$).

a) K 的一个子集 I 是一个分式理想, 必须并且只需存在环 A 的一个非零理想 J 和一个非零的 $d \in \mathrm{A}$, 使得

$$\mathrm{I} = d^{-1}\mathrm{J}.$$

b) 设 I 和 J 是两个分式理想, 又设 (I : J) 是使得 $x\mathrm{J} \subset \mathrm{I}$ 的 $x \in \mathrm{K}$ 的集合. 证明这是一个分式理想 (经常称为从 J 到 I 内的**传递子**).

c) 设 I 和 J 是两个分式理想, 用 $\mathrm{I} + \mathrm{J}$ 表示和 $x + y$ 的集合, 其中的 $x \in \mathrm{I}, y \in \mathrm{J}$, 而 IJ 表示 K 的这样的元素的集合: 它们可以写成乘积 xy 的 (有限) 和的形式, 其中的 $x \in \mathrm{I}, y \in \mathrm{J}$ (对于 $\mathrm{I}, \mathrm{J} \subset \mathrm{A}$ 这种情形, 参见 §8 习题 10). 证明 $\mathrm{I} + \mathrm{J}$, IJ 和 $\mathrm{I} \cap \mathrm{J}$ 是环 A 的分式理想. 推广 §8 习题 10 的 a) 的公式.

d) 称一个分式理想 I 是**可逆的**, 如果存在一个分式理想 J, 使得

$$\mathrm{IJ} = \mathrm{A};$$

证明这时 J 是唯一的, 并且由关系

$$\mathrm{J} = (\mathrm{A} : \mathrm{I})$$

给定 (这时称 J 是 I 的**逆**, 并且记作 I^{-1}). 换句话说, I 是可逆的, 必须并且只需

$$(\mathrm{A} : \mathrm{I})\mathrm{I} = \mathrm{A}.$$

e) 一个理想 I 是可逆的, 必须并且只需存在元素有限个元素 $x_k \in \mathrm{I}$ 和 $y_k \in (\mathrm{A} : \mathrm{I})$, 使得

$$1 = \sum x_k y_k,$$

x_k 组成 A-模 I 的一个生成元组 (因此一个可逆的分式理想是有限生成的 —— 经常说一个分式理想是有限生成的, 如果它作为 A-模是有限生成的).

f) 称 A 是一个 **Dedekind 整环**, 如果 A 的所有分式理想是可逆的. 证明这时 A 的分式理想的集合配备了运算 $(\mathrm{I},\mathrm{J}) \to \mathrm{IJ}$ 构成一个群.

g) 设 A 是一个 Dedekind 整环, 则 A 的所有非零素理想是极大的.

[Dedekind 整环首先在代数数理论中被引进 (参见 §34 的习题 48—50), 而给它以一般的和抽象的定义已经很晚了. 这些环的主要性质是: *在一个 Dedekind 整环内, 所有非零理想都可以写成一个素理想的乘积, 并且如果不考虑素理想的次序, 书写的方式是唯一的* (参见 §18, 习题 7). 并且, 这个性质刻画了 Dedekind 整环的特征. 在实际中更方便使用的是第三个特征将在 §34, 习题 50 给出. 最后注意 Dedekind 整环不仅涉及代数数理论, 而且还涉及代数曲线和代数几何的许多其他问题的研究. 这些都证实了引进 "抽象的" Dedekind 整环的必要性.

作为 Dedekind 整环的最初等的例子, 除了环 $\mathbf{Z}$, 我们还举出环 $\mathbf{Z}[\sqrt{d}]$, 其中 $d \equiv 2$ 或 $3 \pmod 4$).]

¶15. 设 K 是一个环, 而 X 是一个集合. 用

$$\mathrm{K}^{(\mathrm{X})}$$

表示所有这样的映射

$$u : \mathrm{X} \to \mathrm{K}$$

的集合, 其中满足 $u(x) \neq 0$ 的 $x \in \mathrm{X}$ 的数目是*有限的*. 证明 $\mathrm{K}^{(\mathrm{X})}$ 是 (由从 X 到 K 内的所有映射组成的) 左 K-模 K^{X} 的一个子模. 对于每个 $x \in \mathrm{X}$, 考虑 $\mathrm{K}^{(\mathrm{X})}$ 的元素 e_x, 其定义是

$$e_x(y) = \begin{cases} 1, & \text{如果 } y = x, \\ 0, & \text{如果 } y \neq x. \end{cases}$$

证明族 $(e_x)_{x\in\mathrm{X}}$ 是左 K-模 $\mathrm{K}^{(\mathrm{X})}$ 的一个基, 并且关于这个基所有 $u \in \mathrm{K}^{(\mathrm{X})}$ 的分量是标量 $u(x)$, 即对于所有 $u \in \mathrm{K}^{(\mathrm{X})}$ 有

$$u = \sum_{x\in\mathrm{X}} u(x) \cdot e_x.$$

设 f 是从 X 到一个左 K-模 M 的映射, 证明存在唯一的一个同态

$$\bar{f} : \mathrm{K}^{(\mathrm{X})} \to \mathrm{M},$$

使得

$$\bar{f}(e_x) = f(x) \quad \text{对于所有 } x \in \mathrm{X}.$$

(模 $\mathrm{K}^{(\mathrm{X})}$ 的元素一般称为 X 的元素的系数在 K 内的**形式线性组合**, 并且经常把这个模的基的元素 e_x 和对应的元素 $x \in \mathrm{X}$ 等同.)

¶16. 设 K 和 L 是两个交换环, 而 $j_1, \cdots, j_n$ 是从 K 到 L 内的两两不同的同态. 证明 $j_1, \cdots, j_n$ 在从 K 到 L 内的所有映射的 L-模内是线性无关的 (Dedekind 定理).

[写出 $j_1, \cdots, j_n$ 之间的线性关系, 并且利用等式 $j_k(xy) = j_k(x)j_k(y)$ 归结为 $n-1$ 个同态的情形.]

¶17. 设 G 是一个交换群, K 是一个交换环, 而 $\rho_1, \cdots, \rho_n$ 是从 G 到乘法群 K^* (K 的可逆元的群) 的两两不同的同态. 证明 $\rho_1, \cdots, \rho_n$ 是 K 上线性无关的, 即如果 $\alpha_1, \cdots, \alpha_n \in K$ 满足

$$\alpha_1\rho_1(s) + \cdots + \alpha_n\rho_n(s) = 0 \quad 对于所有\ s \in G,$$

则有 $\alpha_1 = \cdots = \alpha_n = 0$ (关于 n 进行归纳推理). 例子: 设 $c_1, \cdots, c_n$ 是两两互异的复数, 考虑 $t \in \mathbf{R}$ 的函数

$$e^{c_1 t}, \cdots, e^{c_n t},$$

证明它们是在 $\mathbf{C}$ 上线性无关的.

§12 线性映射, 矩阵

1. 同态的定义

设 L 和 M 是环 K 上的两个左模. 称从 L 到 M 内的映射

$$f : L \to M$$

为一个**同态**或**线性映射**, 如果有

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y), \quad 任意\ x, y \in L, \lambda, \mu \in K.$$

称所有从 L 到 M 上的双射的同态为从 L 到 M 上的**同构**; 称 L 和 M 是**同构的**, 如果存在从 L 到 M 上的同构.

给定一个左 K-模 M, 称所有从 M 到 M 内的同态为 M 的**自同态** (有时称为 M 上的**线性算子**), 从 M 到自身上的所有同构为**自同构**.

设 L 和 M 是两个左 K-模, 从 L 到 M 内的映射 f 是线性的, 必须且只需有关系

$$f(x + y) = f(x) + f(y), \quad 任意\ x, y \in L,$$
$$f(\lambda x) = \lambda f(x), \quad 任意\ \lambda \in K\, x \in L.$$

在第二个关系里取 $\lambda = 0$ 即得

$$f(0) = 0.$$

此外, 如果 f 是从 L 到 M 内的同态, 则对于任意整数 n, 向量 $x_1, \cdots, x_n$ 和标量 $\lambda_1, \cdots, \lambda_n$ 有关系

$$f(\lambda_1 x_1 + \cdots + \lambda_n x_n) = \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n). \tag{1}$$

这个关系当 $n = 2$ 时归结为同态的定义本身, 在一般情形用关于 n 的归纳法证明:

$$\begin{aligned} f(\lambda_1 x_1 + \cdots + \lambda_n x_n) &= f[(\lambda_1 x_1 + \cdots + \lambda_{n-1} x_{n-1}) + \lambda_n x_n] \\ &= f(\lambda_1 x_1 + \cdots + \lambda_{n-1} x_{n-1}) + f(\lambda_n x_n) \\ &= \lambda_1 f(x_1) + \cdots + \lambda_{n-1} f(x_{n-1}) + f(\lambda_n x_n), \end{aligned}$$

这正是所宣布的结果.

后面我们将经常利用关系 (1), 而一般并不明确注明引用了它.

定理 1　设 $f: \mathrm{L} \to \mathrm{M}$ 和 $g: \mathrm{M} \to \mathrm{N}$ 是模的同态, 则复合映射 $g \circ f$ 还是一个同态, 如果 f 和 g 是同构, 则 $g \circ f$ 是同构. 模的一个同构的逆映射是模的同构.

设 $h = g \circ f$, 则有

$$\begin{aligned} h(\lambda x + \mu y) &= g[f(\lambda x + \mu y)] = g[\lambda f(x) + \mu f(y)] \\ &= \lambda g[f(x)] + \mu g[f(y)] = \lambda h(x) + \mu h(y), \end{aligned}$$

这就表明 h 是一个同态. 如果进一步假设 f 和 g 是同构的, 即是双射的, 则 h 也是双射, 从而它是一个同构.

设 f 是一个同构, 为了证明映射 f^{-1} (它是双射的) 是一个同构, 只需证明它是线性的, 换句话说, 证明对于任意 $\lambda, \mu \in \mathrm{K}$ 和 $u, v \in \mathrm{M}$ 我们有

$$f^{-1}(\lambda u + \mu v) = \lambda f^{-1}(u) + \mu f^{-1}(v);$$

或写成

$$\lambda u + \mu v = f[\lambda f^{-1}(u) + \mu f^{-1}(v)].$$

由于 f 是线性的, 这可以平凡地被验证, 由此定理证毕.

类似 §7, 第 8 小节从定理 1 推出 "X 和 Y 是同构的" 是左 K-模之间的一个等价关系.

在实际中, 经常把同构的模 L 和 M 看作是等同的. 更严格地说, 如果选定从 L 到 M 上的一个同构 f, 那么可以把 L 的元素之间的代数关系翻译成这些元素在 f 下的像之间的类似关系, 因此翻译 L 的所有性质为 M 的类似性质. 读者尽可能地验证这一事实是有裨益的.

定理 2　设 $f: \mathrm{L} \to \mathrm{M}$ 是模的同态. L 的任意一个子模在 f 下的像是 M 的一个子模. M 的一个子模在 f 下的逆像是 L 的一个子模.

设 L' 是 L 的一个子模. 假定 $f(\mathrm{L}')$ 含有 M 的两个元素 u, v; 那么就可以写出 $u = f(x), v = f(y)$, 其中 $x, y \in \mathrm{L}'$. 由于 f 是线性的, 我们有

$$\lambda u + \mu v = f(\lambda x + \mu y) = f(z),$$

由于 L' 是 L 的一个子模, 其中的 $z = \lambda x + \mu y \in \mathrm{L}'$. 于是 $f(\mathrm{L}')$ 含有 $\lambda u + \mu v$, 这里 λ 和 μ 是任意标量, 这就证明了定理的第一个断言. 而第二个断言的证明是类似的.

这里有两个特殊情形, 使得 $f(x) = 0$ 的 $x \in \mathrm{L}$ 的集合称为 f 的**核**, 它是 L 的一个子模, 按照 §7 的第 9 小节记作

$$\mathrm{Ker}(f).$$

而 f 的**像**

$$\mathrm{Im}(f) = f(\mathrm{L})$$

是 M 的一个子模. 我们提醒 (§7, 定理 8) 当且仅当 f 的核缩减为 0, f 是单射的.

2. 从有限生成自由模到任意模内的同态

以下结果是基本的:

定理 3 设 L 是一个有限生成的(*) 自由的左 K-模, $a_1, \cdots, a_p$ 是 L 的一个基, 设 M 是任意一个左 K-模, $c_1, \cdots, c_p$ 是 M 的给定元素. 则存在唯一的一个满足

$$f(a_i) = c_i \quad 对于\ 1 \leqslant i \leqslant p$$

的从 L 到 M 内的同态 f. f 是单射的 (对应的, 满射的), 必须并且只需向量 $c_1, \cdots, c_p$ 是线性无关的 (对应的, 生成 M 的).

考虑到第 1 小节的关系 (1), 同态 f 如果存在, 必然由公式

$$f(\xi_1 a_1 + \cdots + \xi_p a_p) = \xi_1 c_1 + \cdots + \xi_p c_p \tag{2}$$

给定, 这已经表明了 f 的唯一性.

为了确立 f 的存在性, 我们注意到, 对于所有 $x \in \mathrm{L}$, 存在一族标量 ξ_i $(1 \leqslant i \leqslant p)$, 使得

$$x = \xi_1 a_1 + \cdots + \xi_p a_p,$$

即

$$\xi_i = f_i(x) \quad (1 \leqslant i \leqslant p),$$

这里 $f_i : \mathrm{L} \to \mathrm{K}$ 是模 L 关于基 $a_1, \cdots, a_p$ 的坐标函数 (§11, 第 4 小节). 这样一来, 公式 (2) 确实定义了一个从 L 到 M 内的映射 f, 并且根据前面所说的, 它由关系

$$f(x) = f_1(x)c_1 + \cdots + f_p(x)c_p$$

给定. 所有的事情便归结为证明 f 是变换向量 a_i 成 c_i 的一个同态. f 变换向量 a_i 成 c_i 的断言由 §11 第 4 小节所确立的公式,

$$f_i(a_j) = \begin{cases} 1, & 如果\ i = j, \\ 0, & 如果\ i \neq j \end{cases}$$

和关系

$$c_i = 0 \cdot c_1 + \cdots + 0 \cdot c_{i-1} + 1 \cdot c_i + 0 \cdot c_{i+1} + \cdots + 0 \cdot c_p$$

(*) 可以推广定理 3 到任意自由模的情形, 读者容易验证这一事实.

得到. 为了确立 f 是一个同态, 我们首先注意到在 §11 第 4 小节所证明的公式

$$f_i(\lambda x + \mu y) = \lambda f_i(x) + \mu f_i(y),$$

即坐标函数 $f_i : \mathrm{L} \to \mathrm{K}$ 是左 K-模之间的同态. 因此有

$$\begin{aligned} f(\lambda x + \mu y) &= f_1(\lambda x + \mu y)c_1 + \cdots + f_p(\lambda x + \mu y)c_p \\ &= [\lambda f_1(x) + \mu f_1(y)]c_1 + \cdots + [\lambda f_p(x) + \mu f_p(y)]c_p \\ &= \lambda[f_1(x)c_1 + \cdots + f_p(x)c_p] + \mu[f_1(y)c_1 + \cdots + f_p(y)c_p] \\ &= \lambda f(x) + \mu f(y), \end{aligned}$$

这就证明了 f 是线性的.

剩下要证明 f 是满射和单射的条件. 首先, 显然像 $f(\mathrm{L})$ 是向量 $c_1, \cdots, c_p$ 的线性组合的集合, 即 f 映射 L 到由 $c_1, \cdots, c_p$ 生成的 M 的子模上, 因而 f 是满射的, 必须且只需向量 c_i 生成 M.

其次, 为了 f 是单射的, 只需 $f(x) = 0$ 蕴含 $x = 0$, 即关系

$$\xi_1 c_1 + \cdots + \xi_p c_p = 0$$

蕴含

$$\xi_1 a_1 + \cdots + \xi_p a_p = 0,$$

(由于 a_i 是线性无关的) 这就蕴含

$$\xi_1 = 0, \quad \cdots, \quad \xi_p = 0.$$

定理 3 证明完毕.

推论 1 一个左 K-模 M 是有限生成自由的, 必须且只需存在一个整数 n, 使得 M 同构于 K^n. 更精确地说, 设 $c_1, \cdots, c_n$ 是 M 的元素, 而 $e_1, \cdots, e_n$ 是 K^n 的典范基. 那么向量 c_i 组成 M 的一个基, 必须且只需存在从 K^n 到 M 上的一个同构, 把向量 e_i 映射到向量 c_i.

根据定理 3, 总存在唯一的一个同态

$$f : \mathrm{K}^n \to \mathrm{M},$$

使得

$$f(e_i) = c_i \ (1 \leqslant i \leqslant n).$$

c_i 组成 M 的一个基, 必须且只需它们是线性无关的, 并且生成 M, 即 f 是单射的和满射的, 故得推论.

推论 2 一个左 K-模 M 是有限生成的, 必须且只需存在一个整数 n 和一个从 K^n 到 M 内的同态.

更精确地说, 向量 $c_1, \cdots, c_n$ 生成 M, 必须并且只需使得 $f(e_i) = c_i$ 的同态 $f: \mathrm{K}^n \to \mathrm{M}$ 是满射的, 由此得到条件的必要性. 条件是充分的这一断言从 K^n 是有限生成的以及以下观察得到: 设

$$f: \mathrm{L} \to \mathrm{M}$$

是 K-模的同态, 如果 f 是满射的, 并且 L 是有限生成的, 则 M 是有限生成的. 事实上, 设 $a_1, \cdots, a_n$ 是 L 的生成元组, 并且令 $b_i = f(a_i)$ $(1 \leqslant i \leqslant n)$. 对于所有 $y \in \mathrm{M}$, 存在一个 $x \in \mathrm{L}$, 使得 $y = f(x)$. 写出 $x = \xi_1 a_1 + \cdots + \xi_n a_n$ 并且利用第 1 小节的关系 (1), 我们发现 $y = \xi_1 b_1 + \cdots + \xi_n b_n$, 诸 b_i 生成 M, 故 M 是有限生成的.

3. 同态和矩阵

定理 3 让我们能够构造从一个有限生成自由的 K-模 L 到一个任意模 M 的同态. 为了进展得更远, 我们要假定 M 自己也是有限生成的和自由的. 以下我们用 $a_1, \cdots, a_q$ 表示 L 的一个基, 用 $b_1, \cdots, b_p$ 表示 M 的一个基. 此外, 我们假定 L 和 M 是右 K-模. (读者自己考察左模的情形, 不过如果 K 是交换的, 它跟右模的情形没有区别.)

给定一个同态 $f: \mathrm{L} \to \mathrm{M}$. 考虑 L 内的一个向量

$$x = a_1\xi_1 + \cdots + a_q\xi_q \tag{3}$$

和它在 M 内的像

$$f(x) = y = b_1\eta_1 + \cdots + b_p\eta_p. \tag{4}$$

我们打算找到通过 x 关于 L 的 $a_1, \cdots, a_q$ 基的坐标 $\xi_1, \cdots, \xi_q$ 计算 $f(x)$ 关于 M 的 $b_1, \cdots, b_p$ 基的坐标 $\eta_1, \cdots, \eta_p$ 的一个公式.

为此, 令

$$\begin{cases} f(a_1) = c_1 = b_1\alpha_{11} + b_2\alpha_{21} + \cdots + b_p\alpha_{p1}, \\ \qquad\qquad \cdots\cdots\cdots\cdots \\ f(a_q) = c_{1q} = b_1\alpha_{1q} + b_2\alpha_{2q} + \cdots + b_p\alpha_{pq}, \end{cases} \tag{5}$$

其中向量 $f(a_1), \cdots, f(a_q)$ 关于基 $b_1, \cdots, b_p$ 的坐标凸显了出来: 这些坐标 α_{ij} 仅依赖 f 以及在 L 和 M 内所选定的两个基. 于是有

$$\begin{aligned} f(x) &= c_1\xi_1 + \cdots + c_q\xi_q \\ &= (b_1\alpha_{11} + b_2\alpha_{21} + \cdots + b_p\alpha_{p1})\xi_1 + \cdots + (b_1\alpha_{1q} + b_2\alpha_{2q} + \cdots + b_p\alpha_{pq})\xi_q, \end{aligned}$$

把第二个等号右端的加法按列实施, 就得到对于 $f(x)$ 的坐标要找的值, 即

$$\begin{cases} \eta_1 = \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \cdots + \alpha_{1q}\xi_q, \\ \qquad\cdots\cdots\cdots\cdots \\ \eta_p = \alpha_{p1}\xi_1 + \alpha_{p2}\xi_2 + \cdots + \alpha_{pq}\xi_q. \end{cases} \tag{6}$$

这个公式称为 f **关于 L 的基** $(a_j)_{1\leqslant j\leqslant q}$ **和 M 的基** $(b_i)_{1\leqslant i\leqslant p}$ **的方程**, 可以把它写成紧凑的形式:

$$\eta_i = \sum_{i=1}^{q} \alpha_{ij}\xi_{ij} \quad (1 \leqslant i \leqslant p). \tag{6'}$$

反之, 我们给定标量 $\alpha_{ij} \in \mathrm{K}(1 \leqslant i \leqslant p, 1 \leqslant j \leqslant q)$, 并且用公式 (6) 定义一个从 L 到 M 内的映射 f, f 把关于 L 的给定的基坐标为 $\xi_1, \cdots, \xi_q$ 的向量 $x \in \mathrm{L}$ 变换为关于 M 的给定的基坐标为按照关系 (6) 算出的 $\eta_1, \cdots, \eta_p$ 的向量. 那么 f 是从 L 到 M 内的一个同态. 事实上, 我们有

$$\begin{aligned} f(x) &= b_1\eta_1 + \cdots + b_p\eta_p \\ &= b_1(\alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \cdots + \alpha_{1q}\xi_q) + \cdots + b_p(\alpha_{p1}\xi_1 + \alpha_{p2}\xi_2 + \cdots + \alpha_{pq}\xi_p) \\ &= c_1\xi_1 + \cdots + c_q\xi_q, \end{aligned}$$

其中向量 c_j 由公式 (5) 给定, 我们发现 f 是一个同态, 其存在性由定理 3 保证.

最后注意给定同态 f 后, 存在唯一的一组标量 α_{ij}, 使得 f 由关系 (6) 给定, 因为刚做的计算表明 α_{ij} 必然是向量 $f(a_i)$ 关于 M 的基 $(b_j)_{1\leqslant j\leqslant q}$ 的坐标, 这就完全确定了 α_{ij}.

现在回到公式 (6), 知道了它就能够定义同态 f. 为了记住它, 显然只需知道由常量 α_{ij} (称为出现在公式 (6) 里的**系数**) 组成的表格

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1q} \\ \vdots & \vdots & & \vdots \\ \alpha_{p1} & \alpha_{p2} & \cdots & \alpha_{pq} \end{pmatrix}, \tag{7}$$

形如 (7) 的表格称为 p **行** q **列的元素在环 K 内的矩阵** (α_{ij} 也称为所提到的矩阵的**元素**), 我们说 (7) 是**同态** f **关于 L 的基** $(a_j)_{1\leqslant j\leqslant q}$ **和 M 的基** $(b_i)_{1\leqslant i\leqslant p}$ **的矩阵**. 矩阵的概念对于同态起的作用类似于坐标概念对于向量起的作用.

当不做明晰的计算时, 经常把矩阵 (7) 写成紧凑的形式

$$(\alpha_{ij})_{1\leqslant i\leqslant p, 1\leqslant j\leqslant q},$$

这种写法还有一个好处是表明一个矩阵其实就是 K 的元素的以正整数序偶 (i, j) 为指标的元素的一个族, 其中 $1 \leqslant i \leqslant p, 1 \leqslant j \leqslant q$.

当 $\mathrm{L}=\mathrm{M}$, 即当 f 是模 L 的一个自同态时, 经常在 L 内和 M 内使用同一个基 $a_1,\cdots,a_p$, 这样就可以谈论**关于 L 的一个基 L 的自同态的矩阵**. 所述的矩阵显然有 p 行 p 列, 人们说这是**一个元素在 K 内的 p 阶方阵**.

注 1 给定有限生成自由的 K-模的一个同态 $f:\mathrm{L}\to\mathrm{M}$, 不能笼统谈论 f 的矩阵; 为了使这个概念有意义, 应当首先选择 L 的一个基和 M 的一个基, 而所得到的矩阵显然依赖两个基的选择 (在 §15, 将看到当在 L 和 M 中的基改变时矩阵会发生什么变化).

然而, 当 $\mathrm{L}=\mathrm{K}^q$ 和 $\mathrm{M}=\mathrm{K}^p$ 时, 我们指定选择 L 的典范基和 M 的典范基 (§11, 例 12). 给定一个同态

$$f:\mathrm{K}^q\to\mathrm{K}^p,$$

那么谈论 f 的矩阵是合理的 (约定:关于 K^p 的典范基和 K^q 的典范基). 如果用 $(\alpha_{ij})_{1\leqslant i\leqslant p,1\leqslant j\leqslant q}$ 表示这个矩阵, 那么 f 正是这样的映射: 它变换每个向量 $x=(\xi_1,\cdots,\xi_q)\in\mathrm{K}^q$ 为由 (6) 给定其分量的向量

$$f(x)=(\eta_1,\cdots,\eta_p)\in\mathrm{K}^p.$$

反之, 同样的构造使得有可能令每个矩阵 $(\alpha_{ij})_{1\leqslant i\leqslant p,1\leqslant j\leqslant q}$ 对应一个从 K^p 到 K^q 内的同态, 即其矩阵是给定矩阵的同态.

这个考虑表明存在从 K^p 到 K^q 内的同态的集合到 p 行 q 列的 (元素在 K 内的) 矩阵的集合上的一个 "典范" 双射. 经常会用到这个对应, 而不明确注明.

注 2 设 L 和 M 是有限生成自由 K-模, $(a_j)_{1\leqslant j\leqslant q}$ 是 L 的一个基, 而 $(b_i)_{1\leqslant i\leqslant p}$ 是 M 的一个基. 引进同构

$$u:\ \mathrm{K}^q\to\mathrm{L},\quad v:\mathrm{K}^p\to\mathrm{M},$$

它们分别映射 K^p 和 K^q 的基为的 L 和 M 给定的基 (定理 3 的推论 1).

给定一个同态 $f:\mathrm{L}\to\mathrm{M}$; 借助 u 和 v, 演绎出 (定理 1) 一个同态

$$\bar{f}=v^{-1}\circ f\circ u:\mathrm{K}^q\to\mathrm{K}^p;$$

那么, f 关于 L 的基 (a_i) 和 M 的基 (b_j) 的矩阵就等于 $\bar{f}$ 的 (关于 K^q 和 K^p 的典范基) 的矩阵. 事实上, 我们考虑向量

$$x=a_1\xi_1+\cdots+a_p\xi_p,\quad f(x)=b_1\eta_1+\cdots+b_q\eta_q,$$

其中的 η_j 通过 ξ_i 由前面的 (6) 给定. 由 u 的构造得到

$$x=u(\xi_1,\cdots\xi_p),$$

同样有

$$f(x) = v(\eta_1, \cdots, \eta_q);$$

于是

$$(\eta_1, \cdots, \eta_p) = v^{-1}[f(x)] = v^{-1}[f(u(\xi_1, \cdots \xi_q))] = \bar{f}(\xi_1, \cdots \xi_q);$$

因此公式 (6) 也是 $\bar{f}$ (关于典范基) 的方程, 我们的断言得证.

我们发现还可以这样引进矩阵, 首先证明所有的从 K^q 到 K^p 内的同态由形如 (6) 的关系给定, 其次注意到下列事实: 如果有一个有限生成的自由的模的同态 $f : L \to M$, 那么 L 的一个基和 M 的一个基的选取使得 L 等同于 K^q, M 等同于 K^p, 并且因此 f 等同于从 K^q 到 K^p 内的由 (6) 给定的同态 $\bar{f}$, 并且以典范的方式由一个矩阵表示. 我们重新发现了一旦选定了 L 和 M 的基之后, 把一个矩阵依附于一个同态 f 的可能性.

4. 同态和矩阵的例子

我们要列举几个同态和用矩阵表示同态的例子.

例 1 设 L 是由给定起点 O 的通常平面上的向量组成的实向量空间, 考虑映射 $f : L \to L$, 它变换所有起点为 O 的向量为绕 O 的一个角为 θ (给定) 的旋转得到的向量. 这显然是一个线性映射 (因为旋转把平行四边形变换为一个平行四边形, 并且中心为 O 的旋转跟中心为 O 的**位似**变换可交换). 设 Ox, Oy 为平面上的直角坐标轴, i(*) 和 j 是 Ox 和 Oy 上的单位向量. 几何的推理显然证明有关系

$$\begin{aligned} f(i) &= i \cdot \cos\theta + j \cdot \sin\theta, \\ f(j) &= -i \cdot \sin\theta + j \cdot \cos\theta, \end{aligned}$$

因此, f 关于基 (i, j) 的矩阵就是

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

(不要忘记一个同态的矩阵 (7) 的行由关系 (5) 中的列得到, 这就解释了这里求得的结果).

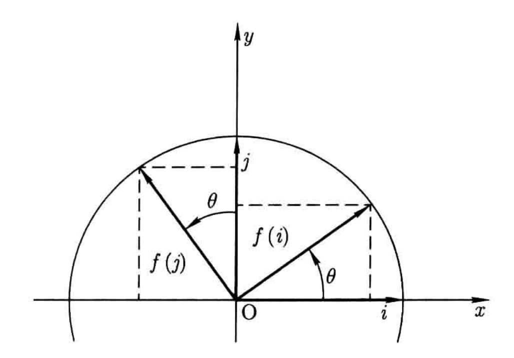
例 2 取 $L = M$, 并且假定 K 是交换的, 那么对于所有 $\lambda \in K$, 映射 $x \to \lambda x$ (L 内的比为 λ 的位似) 是线性的. 由于它变换 L 的一个基 $a_1, \cdots, a_n$ 为

$$\lambda \cdot a_1 = \lambda \cdot a_1 + 0 \cdot a_2 + \cdots + 0 \cdot a_n,$$

$$\cdots\cdots\cdots\cdots$$

$$\lambda \cdot a_n = 0 \cdot a_1 + 0 \cdot a_2 + \cdots + \lambda \cdot a_n,$$

(*) 向量 i 跟复数 i 毫无关系.



我们发现比为 λ 的位似对于 L 的一个任意的基的矩阵是

$$\begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix},$$

我们说这是其元素在 L 内的标量矩阵.

我们注意这里 f 的矩阵不依赖 L 的基 $a_1, \cdots, a_n$ 的选择. 可以容易地证明位似是 L 的唯一的具有这一性质的同态.

例 3 称所有从右 K-模 L 到右 K-模 K 内的同态为右 K-模 L 上的**线性型**或**余向量**, 换句话说, 一个线性型是一个映射

$$f : \mathrm{L} \to \mathrm{K},$$

使得有

$$f(x\lambda + y\mu) = f(x)\lambda + f(y)\mu, \quad 任意\ \lambda, \mu \in \mathrm{K}, x, y \in \mathrm{L}.$$

例如, §11 第 4 小节表明 L 关于 L 的一个基的坐标函数是 L 上的线性型.

设 f 是 L 上的一个线性型, 并且选择 (如果可能) L 的一个基 $a_1, \cdots, a_q$. 为了像对于所有有限生成自由模的同态那样, 让 f 对应一个矩阵, 还必须选择右 K-模 K 的一个基: 我们选择它的一个由向量 1 组成的典范基. 为了计算 f 关于这两个基的矩阵, 应当写出

$$\begin{aligned} &f(a_1) = 1 \cdot \alpha_{11}, \\ &\cdots\cdots\cdots\cdots \\ &f(a_q) = 1 \cdot \alpha_{1q}; \end{aligned}$$

显然更简单地, 令

$$f(a_1) = \alpha_1,$$
$$\cdots\cdots\cdots\cdots$$
$$f(a_q) = \alpha_q,$$

那么, f 通过所选择的 L 和 K 的基就表示成**行矩阵**

$$(\alpha_1, \cdots, \alpha_q).$$

由于有公式

$$f(a_1\xi_1 + \cdots + a_q\xi_q) = \alpha_1\xi_1 + \cdots + \alpha_p\xi_p,$$

该行矩阵确定了 f. 经常称 α_i 是 f **关于 L 的基** $(a_i)_{1\leqslant i\leqslant q}$ **的系数**.

在特殊情形 $\mathrm{L} = \mathrm{K}^q$ 时, 自然要求选择 L 的典范基, 就简单地称 α_i 是 f (关于 K^q 的典范基) 的**系数**.

例 4 设 M 是右 K-模, 考虑右 K-模的一个同态

$$f : \mathrm{K} \to \mathrm{M}.$$

令

$$f(1) = c \in \mathrm{M},$$

则有

$$f(\xi) = f(1 \cdot \xi) = f(1)\xi = c\xi,$$

于是知道了 c 就确定了 f (在实际中, 对于同态 f 和 M 的元素 c 不加区别). 假定 M 具有一个基 $b_1, \cdots, b_p$, 在 K 内使用典范基, 令

$$c = f(1) = b_1\alpha_1 + \cdots + b_p\alpha_p,$$

我们发现 f 关于所考虑的 K 和 M 的基的矩阵是由向量 c 关于 M 的所选定的基的分量组成的**列矩阵**

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix}.$$

后面有时会把 M 的每个向量 (一旦选定 M 的一个基后) 等同于列矩阵, 它的元素是所考虑的向量关于 M 内所选定的基的坐标. 这个观点的合理性在后面 (§14, 第 4 小节) 会得到验证.

前面的例子都是纯代数性质的 (由于它们都是关于有限生成的自由模的). 反之分析提供同态的例子, 企图用矩阵 (即使是无限的矩阵) 表示它们是完全人为的. 这里是几个例子.

例 5 设 a 和 b 是两个实数, $a < b$, 用 X 表示满足 $a \leqslant t \leqslant b$ 的 $t \in \mathbf{R}$ 的集合 (X 其实就是区间 $[a, b]$), 用 L 表示由处处连续的映射 $f : \mathrm{X} \to \mathbf{R}$ 组成的实向量空间 (§10, 例 4 和 7). 设 $\mathrm{N}(s, t)$ 是在正方形 $\mathrm{X} \times \mathrm{X}$ 上定义并且连续的函数. (利用函数 N 在正方形 $\mathrm{X} \times \mathrm{X}$ 上一致连续这个事实) 可以证明对于所有函数 $f \in \mathrm{L}$, 函数

$$f^*(s) = \int_a^b \mathrm{N}(s,t)f(t)dt \quad (a \leqslant s \leqslant b)$$

在 X 上仍然连续, 即 $f^* \in \mathrm{L}$. 这一点交代清楚了, 就可以说 L 到 L 内的映射 $u : f \to f^*$ 是线性的. 事实上, 令 $f + g = h$, 则有

$$\begin{aligned} h^*(s) &= \int_a^b \mathrm{N}(s,t)h(t)dt = \int_a^b \mathrm{N}(s,t)[f(t)+g(t)]dt \\ &= \int_a^b \mathrm{N}(s,t)f(t)dt + \int_a^b \mathrm{N}(s,t)g(t)dt = f^*(s) + g^*(s), \end{aligned}$$

即 $h^* = f^* + g^*$, 从而有 $u(f+g) = u(f) + u(g)$. 同样证明对于任意 $\lambda \in \mathbf{R}$ 有关系 $u(\lambda f) = \lambda u(f)$.

从 L 到 L 内的映射 u 称为**积分算子**, 这些算子的研究 (尤其是 Hilbert 和 F. Riesz) 导致今日的*泛函分析*在 20 世纪前四分之一创立. 不言而喻, 本书纯代数的并且本质上是平凡的考虑, 对于泛函分析没有任何重大的帮助 —— 否则, 那也太容易了 —— 不过为它提供了合理的术语和应当在什么方向进行探索的粗略思路. 事实上, 在泛函分析中为了得到非平凡的结果所碰到的困难鲜有代数性质的, 最经常是 “分析” 性质的, 并且为了克服这些困难强求使用 “拓扑” 方法 (即建立在 “连续性” 概念基础上的方法). 此外有趣的是, 注意到初等线性代数的发展极大地受到泛函分析发展的影响, 我们更期待尽快看到反向的影响出现.

例 6 取上例中的同一个向量空间 L, 那么从 L 到 $\mathbf{R}$ 内的映射

$$f \to \int_a^b f(t)dt = \mathrm{I}(f)$$

是 L 上的*线性型*. 事实上, 如果 f 和 g 是区间 $[a, b]$ 上的连续函数, 那么函数 $f+g$ 的积分等于 f 和 g 的积分的和; 如果函数 f 乘以一个常数 $\lambda \in \mathbf{R}$, 它的积分也乘以 λ.

例 7 基础环仍是 $\mathbf{R}$, 我们考虑这样两个实向量空间 L 和 M: L 的元素是具有*连续的二阶导数* f'' 的映射 $f : \mathbf{R} \to \mathbf{R}$; 而 M 的元素是所有连续函数 $g : \mathbf{R} \to \mathbf{R}$ (不加任何可导性条件). 当然在 L 和 M 内的线性运算像在 §10 例 4 中那样定义.

选定函数 $a, b, c \in \mathrm{M}$ (即一个变量的连续函数), 对于所有函数 $f \in \mathrm{L}$ 构造函数

$$f^*(t) = a(t)f(t) + b(t)f'(t) + c(t)f''(t),$$

显然 f^* 属于 M, 故得到一个由对于所有 $f \in \mathrm{L}$ 令 $\mathrm{D}(f) = f^*$ 给定的一个映射 $\mathrm{D}: \mathrm{L} \to \mathrm{M}$. 这样定义的 D 是一个同态. 事实上,

$$\begin{aligned}\mathrm{D}(f+g) &= a(f+g) + b(f+g)' + c(f+g)'' \\ &= af + bf' + cf'' + ag + bg' + cg'' = \mathrm{D}(f) + \mathrm{D}(g),\end{aligned}$$

同样证明 $\mathrm{D}(\lambda f) = \lambda \mathrm{D}(f)$.

这类同态涉及线性微分方程理论.

注意同样容易构造向量空间 L 上的线性型. 例如从 L 到 $\mathbf{R}$ 内的映射

$$f \to f''(0)$$

就是这种情形, 它令每一个 $f \in \mathrm{L}$ 对应它在 $t = 0$ 的二阶导数.

(本节习题在 §14 之后.)

§13 同态和矩阵的加法

1. 加法群 Hom(L,M)

设 L 和 M 是两个在任意环 K 上的 (比如, 左) K-模. 用

$$\mathrm{Hom}(\mathrm{L},\mathrm{M}) \quad 或 \quad \mathscr{L}(\mathrm{L},\mathrm{M})$$

表示所有从 L 到 M 内的线性映射的集合. (为避免基础环混淆) 还可能应用记号

$$\mathrm{Hom}_{\mathrm{K}}(\mathrm{L},\mathrm{M}) \quad 或 \quad \mathscr{L}_{\mathrm{K}}(\mathrm{L},\mathrm{M}).$$

定理 1 设 L 和 M 是两个左 K-模. 如果

$$f, g: \mathrm{L} \to \mathrm{M}$$

是从 L 到 M 内的同态, 则映射

$$f + g: x \to f(x) + g(x)$$

也是从 L 到 M 内的同态. 配备了运算 $(f,g) \to f+g$ 的集合 Hom(L,M) 是一个交换群.

令 $h = f + g$. 那么

$$\begin{aligned}h(\lambda x + \mu y) &= f(\lambda x + \mu y) + g(\lambda x + \mu y) \\ &= \lambda f(x) + \mu f(y) + \lambda g(x) + \mu g(y) \\ &= \lambda[f(x) + g(x)] + \mu[f(y) + g(y)] \\ &= \lambda h(x) + \mu h(y),\end{aligned}$$

这就证明了定理的第一个断言.

为了证明第二个断言, 考虑所有从 L 到 M 内的 (线性或非线性) 映射的集合 E, 配备了运算 $(f,g) \to f+g$, 这是一个交换群 (§10, 例 4; 不涉及 L 是否是一个模, 简单地把它看作一个集合). 剩下要做的是证明 Hom(L, M) 是 E 的一个子群. 而 Hom(L, M) 显然含有 E 的中性元 (即从 L 到 M 内的处处取零值的映射), 又如果 f, g 是同态, 则可以像证明第一个断言那样证明 $f-g$ 也是同态, 故得所要求的结果.

定理 1 允许把 Hom(L, M) 称为从 L 到 M 内的**同态群**.

当基础环 K 是交换的时, 甚至可以把 Hom(L, M) 看作一个新的左 K-模. 首先重新取所有从 L 到 M 内的 (线性或非线性) 映射的集合 E; §10 例 4 不仅允许把 E 看作一个加法群, 而且看作一个左 K-模, 一个标量 λ 和一个映射 $f: \mathrm{L} \to \mathrm{M}$ 的乘积 λf 的定义是从 L 到 M 内的映射

$$x \to \lambda f(x)$$

(这里不必假定 K 是交换的). 而当 K 是交换环时, 集合 Hom(L, M) 不仅是 E 的一个子群, 而且是 E 的一个子模. 换句话说, 如果 f 是从 L 到 M 内的一个同态, 则 $f' = \lambda f$ 也是同态. 事实上, 我们有

$$\begin{aligned} f'(\alpha x+\beta y) &= \lambda f(\alpha x+\beta y) = \lambda\alpha \cdot f(x) + \lambda\beta \cdot f(y) \\ &= \alpha\lambda \cdot f(x) + \beta\lambda \cdot f(y) = \alpha f'(x) + \beta f'(y), \end{aligned}$$

这正如所断言的那样. 于是确实可以把 Hom(L, M) 看作一个左 K-模.

举例说, 如果 L 和 M 是实 (对应的, 复) 向量空间, 那么可以把 Hom(L, M) 看作一个实 (对应的, 复) 向量空间.

2. 矩阵的加法

在前面的内容中, 现在假定 L 和 M 是有限生成的自由右 K-模. 选择 L 的一个基 $(a_j)_{1\leqslant j\leqslant q}$ 和 M 的一个基 $(b_i)_{1\leqslant i\leqslant p}$. 给定从 L 到 M 内的两个同态 f 和 g, 它们关于所考虑的 L 和 M 的基的矩阵是

$$\mathrm{A} = (\alpha_{ij})_{1\leqslant i\leqslant p, 1\leqslant j\leqslant q},$$
$$\mathrm{B} = (\beta_{ij})_{1\leqslant i\leqslant p, 1\leqslant j\leqslant q},$$

于是有

$$f(a_j) = b_1\alpha_{1j} + \cdots + b_p\alpha_{pj},$$
$$g(a_j) = b_1\beta_{1j} + \cdots + b_p\beta_{pj}.$$

令 $h = f+g$, h 关于所考虑的基的矩阵表示为

$$\mathrm{C} = (\gamma_{ij})_{1\leqslant i\leqslant p, 1\leqslant j\leqslant q},$$

我们有

$$h(a_j) = f(a_j) + g(a_j) = b_1(\alpha_{1j} + \beta_{1j}) + \cdots + b_p(\alpha_{pj} + \beta_{pj}),$$

因此 C 的元素是

$$\gamma_{ij} = \alpha_{ij} + \beta_{ij} \quad (1 \leqslant i \leqslant p, 1 \leqslant j \leqslant q). \tag{1}$$

给定系数在 K 内的两个矩阵 $A = (a_{ij})$ 和 $B = (\beta_{ij})$, 此式引导我们称由关系 (1) 给定的矩阵 $C = (\gamma_{ij})$ 为两个矩阵 A 和 B 的和, 用记号

$$A + B$$

表示它.

我们注意到两个矩阵的和仅当它们的行数相同并且列数相同时才有意义.

如果把矩阵 $(\alpha_{ij})_{1\leqslant i\leqslant p, 1\leqslant j\leqslant q}$ 和 K^{pq} 的元素等同, 那么显然矩阵的加法归结为 K^{pq} 的元素的加法. 因此, 配备了运算 $(A, B) \to A + B$ 的 p 行 q 列的矩阵的集合是一个交换群.

我们有下列结果, 这正是引进矩阵加法运算的目的:

定理 2 设 L 和 M 是有限生成的自由 K-模, f 和 g 是从 L 到 M 内的两个同态. 设 A 和 B 是 f 和 g 关于 L 的一个基 (a_i) 和 M 的一个基 (b_j) 的矩阵. 则 $f + g$ (关于这两个基) 的矩阵是 $A + B$.

最后我们注意不仅能把元素在 K 中的 p 行 q 列的矩阵的集合看作一个加法群, 而且可以看作一个左 (或右) K-模, 只要对于矩阵

$$A = (a_{ij})$$

和 $\lambda \in K$. 用以下公式定义 λA 和 $A\lambda$:

$$\lambda A = (\lambda a_{ij}), \quad A\lambda = (a_{ij}\lambda).$$

(本节习题在 §14 之后.)

§14 矩阵的乘积

1. 模的自同态环

首先证明下列结果:

定理 1 设 L, M, N 是三个模, 给定同态

$$f, g : L \to M \quad 和 \quad h : M \to N,$$

则有关系

$$h \circ (f+g) = h \circ f + h \circ g;$$

给定同态

$$f : \mathrm{L} \to \mathrm{M} \quad \text{和} \quad g, h : \mathrm{M} \to \mathrm{N},$$

则有关系

$$(g+h) \circ f = g \circ f + h \circ f.$$

以证明第一个结果为例. 令 $u = f + g$, 则有

$$h \circ u(x) = h[u(x)] = h[f(x) + g(x)] = h[f(x)] + h[g(x)],$$

这就表明映射 $h \circ u$ 是映射 $h \circ f$ 和 $h \circ g$ 的和, 故得第一个结果, 类似地证明第二个结果.

推论 设 L 是一个环上的模, 则模 L 的同态的集合 Hom(L, L) 配备了运算

$$(f, g) \to f + g, \quad (f, g) \to f \circ g$$

是一个环.

配备了加法的 Hom(L, L) 是一个交换群这一事实从 §13 定理 1 得到. 乘法的结合性从 §2 定理 2 得到, 而中性元的存在性从恒等映射 j_{L} 属于 Hom(L, L) 得到. 最后, 定理 1 表明 "分配律" 的条件满足, 这就完成了证明.

配备了所提到的两个运算的集合 Hom(L, L) 称为**模 L 的同态环**. 一般它不是交换的 (即使基础环是交换的), 后面就会看到这方面的例子.

2. 两个矩阵的乘积

设 L, M, N 是三个有限生成的自由的右模, 并且选择这些模的基分别为 $(a_1, \cdots, a_r)$, $(b_1, \cdots, b_q)$ 和 $(c_1, \cdots, c_p)$. 考虑同态

$$f : \mathrm{M} \to \mathrm{N} \quad \text{和} \quad g : \mathrm{L} \to \mathrm{M}$$

以及复合同态

$$h = f \circ g : \mathrm{L} \to \mathrm{N}.$$

f 关于基 (b_j) 和 (c_k) 的矩阵用

$$\mathrm{A} = (\alpha_{ij})_{1 \leqslant i \leqslant p, 1 \leqslant j \leqslant q}$$

表示, 而 g 关于基 (a_i) 和 (b_j) 的矩阵用

$$\mathrm{B} = (\beta_{jk})_{1 \leqslant j \leqslant q, 1 \leqslant k \leqslant r}$$

表示, 还有 h 关于基 (a_i) 和 (c_k) 的矩阵用

$$\mathrm{C} = (\gamma_{ik})_{1\leqslant i\leqslant p, 1\leqslant k\leqslant r}$$

表示. 我们的任务是通过 A 和 B 计算 C.

设

$$x = a_1\xi_1 + \cdots + a_r\xi_r,$$

是 L 的一个元素, 令

$$g(x) = y = b_1\eta_1 + \cdots + b_q\eta_q,$$
$$h(x) = f(y) = c_1\zeta_1 + \cdots + c_p\zeta_p.$$

由于我们知道 f 和 g 的矩阵 A 和 B, §12 第 3 小节的公式 (6) 表明我们有

$$\zeta_i = \alpha_{i1}\eta_1 + \cdots + \alpha_{iq}\eta_q\ (1 \leqslant i \leqslant p),$$
$$\eta_j = \beta_{j1}\xi_1 + \cdots + \beta_{jr}\xi_r\ (1 \leqslant j \leqslant q),$$

因此就有

$$\zeta_i = \alpha_{i1}(\beta_{11}\xi_1 + \cdots + \beta_{1r}\xi_r) + \alpha_{i2}(\beta_{21}\xi_1 + \cdots + \beta_{2r}\xi_r) + \cdots + \alpha_{iq}(\beta_{q1}\xi_1 + \cdots + \beta_{qr}\xi_r).$$

而 h 的矩阵 $\mathrm{C} = (\gamma_{ik})$ 还由公式

$$\zeta_i = \gamma_{i1}\xi_1 + \cdots + \gamma_{ir}\xi_r$$

给定, 故得

$$\gamma_{i1} = \alpha_{i1}\beta_{11} + \alpha_{i2}\beta_{21} + \cdots + \alpha_{iq}\beta_{q1},$$
$$\cdots\cdots\cdots\cdots$$
$$\gamma_{ir} = \alpha_{i1}\beta_{1r} + \alpha_{i2}\beta_{2r} + \cdots + \alpha_{iq}\beta_{qr},$$

或写成紧凑的形式

$$\gamma_{ik} = \alpha_{i1}\beta_{1k} + \alpha_{i2}\beta_{2k} + \cdots + \alpha_{iq}\beta_{qk} = \sum_{j=1}^{q}\alpha_{ij}\beta_{jk}. \tag{1}$$

这个结果引导我们引进以下定义: 给定两个元素在 K 内的矩阵

$$\mathrm{A} = (\alpha_{ij})_{1\leqslant i\leqslant p, 1\leqslant j\leqslant q}, \quad \mathrm{B} = (\beta_{jk})_{1\leqslant j\leqslant q, 1\leqslant k\leqslant r},$$

称矩阵

$$\mathrm{AB} = (\gamma_{ik})_{1\leqslant i\leqslant p, 1\leqslant k\leqslant r}$$

为 A **和** B **的乘积**, 其中 γ_{ik} 由 (1) 给定.

要注意乘积 AB 仅当 A 的列数等于 B 的行数时才有定义, 此时 AB 跟 A 有同样多的行, 跟 B 有同样多的列.

显然根据前面的定义, 我们可以陈述以下结果:

定理 2 设 L, M, N 是三个有限生成的自由的右 K-模, 而 $(a_k), (b_j), (c_i)$ 分别是 L, M, N 的基, 并且设 $f: \mathrm{M} \to \mathrm{N}$ 和 $g: \mathrm{L} \to \mathrm{M}$ 是同态. 设 f 关于基 (b_j) 和 (c_i) 的矩阵是 A, 而 g 关于基 (a_k) 和 (b_j) 的矩阵是 B, 则 $f \circ g$ 关于基 (a_k) 和 (c_i) 的矩阵是 AB.

现在举几个矩阵乘法的例子.

例 1 令

$$\mathrm{A} = (\alpha_1, \cdots, \alpha_q), \quad B = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_q \end{pmatrix},$$

乘积 AB 是个一行一列的矩阵, 其形式是 (γ). 关系 (1) 显然给出

$$\gamma = \alpha_1\beta_1 + \cdots + \alpha_q\beta_q,$$

基于显然的理由我们说标量 γ 是 "行" A 乘以 "列" B 的乘积.

这个结果帮助我们容易记住矩阵乘法的一般规则. 事实上, 结合我们刚做的约定, 公式 (1) 可以写成

$$\gamma_{ik} = (\alpha_{i1}, \cdots, \alpha_{iq}) \begin{pmatrix} \beta_{1k} \\ \vdots \\ \beta_{qk} \end{pmatrix}.$$

换句话说, AB 的位于第 i 行的元素由 A 的第 i 行乘以 B 的各列而得到. 在实际中我们总是使用这个规则.

例 2 二阶方阵的乘法由公式

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \begin{pmatrix} a'' & b'' \\ c'' & d'' \end{pmatrix} = \begin{pmatrix} a'a'' + b'c'' & a'b'' + b'd'' \\ c'a'' + d'c'' & c'b'' + d'd'' \end{pmatrix}$$

定义. 例如, 如果 x 和 y 是实数, 则有

$$\begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix} \begin{pmatrix} \cos y & \sin y \\ -\sin y & \cos y \end{pmatrix} = \begin{pmatrix} \cos(x+y) & \sin(x+y) \\ -\sin(x+y) & \cos(x+y) \end{pmatrix}.$$

考虑到 §12 例 1, 这个结果表明在平面上复合绕一个点 O 的角为 x 和 y 的旋转得到绕 O 的角为 $x+y$ 的旋转.

例 3 对于所有的环 K 和所有整数 $n \geqslant 1$, 考虑 n 行 n 列矩阵

$$1_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

称为 n **阶单位矩阵**. 这个术语的合理性由以下事实说明, 对于元素在 K 内的任意矩阵 X 和 Y 有关系

$$1_n \mathrm{X} = \mathrm{X}, \quad \mathrm{Y} 1_n = \mathrm{Y}$$

(只要它有意义, 即 X 有 n 行, 而 Y 有 n 列). 根据公式 (1) 这些关系容易得到, 并且可以做如下的几何解释. 设 L 是一个有限生成的自由的右 K-模, 具有由 n 个向量组成的一个基 $(a_i)_{1\leqslant i\leqslant n}$ (比如, 可以取 $\mathrm{L}=\mathrm{K}^n$, 而基为典范基), 那么 §12 第 3 小节的公式 (6) 和这里的事实

$$\alpha_{ij} = \begin{cases} 1, & 如果\ i=j, \\ 0, & 如果\ i\neq j \end{cases}$$

表明关于基 (a_i) 以 1_n 为矩阵的自同态 $j:\mathrm{L}\to\mathrm{L}$ 正是从 L 到 L 上的恒等映射. 为了证明关系 $1_n\mathrm{X}=\mathrm{X}$, 再引进第二个模 M, M 的基 (b_j) 和同态 $f:\mathrm{M}\to\mathrm{L}$, f 关于 L 和 M 所考虑的基 (a_i) 和 (b_j) 有矩阵 X; 定理 1 表明 $1_n\mathrm{X}$ 是同态 $j\circ f$ 关于这些基的矩阵. 由于 $j=j_{\mathrm{L}}$, 我们有 $j\circ f=f$, 故得要证的关系.

3. 矩阵环

我们在本节和前一节所定义的矩阵的加法和乘法遵守通常的计算规则, 只要相应的运算有意义. 对于加法我们已经知道是这样 (§13, 第 2 小节). 至于乘法, 有结合律

$$\mathrm{A(BC)} = \mathrm{(AB)C},$$

只要两端有定义. 为了确认这一事实, 引进同态

$$f:\mathrm{K}^r\to\mathrm{K}^s, \quad g:\mathrm{K}^q\to\mathrm{K}^r, \quad h:\mathrm{K}^p\to\mathrm{K}^q,$$

它们 (关于典范基) 的矩阵分别是 A, B, C, 那么根据定理 2, 矩阵 A(BC) 表示 $f\circ(g\circ h)$, 而矩阵 (AB)C 表示 $(f\circ g)\circ h$, 故得所宣布的结果.

最后还有分配律

$$\mathrm{A(B+C)} = \mathrm{AB}+\mathrm{AC}, \quad \mathrm{(A+B)C} = \mathrm{AC}+\mathrm{BC},$$

跟结合律一样, 用模的同态代替 A, B, C, 并且利用定理 1.

在特殊情形下, 对于所有整数 $n \geqslant 1$, 所有环 K, 用记号

$$\mathrm{M}_n(\mathrm{K})$$

表示元素在 K 内的 n 阶 (即 n 行 n 列) 方阵的集合. 对于这个集合定义运算

$$(\mathrm{A}, \mathrm{B}) \to \mathrm{A} + \mathrm{B}, \quad (\mathrm{A}, \mathrm{B}) \to \mathrm{AB},$$

那么 $\mathrm{M}_n(\mathrm{K})$ 就成为一个环. 事实上, 我们已经知道, $\mathrm{M}_n(\mathrm{K})$ 关于加法是一个交换群. 此外, 刚证明了乘法是结合的, 并且显然 (上面的例 3) 它具有一个乘法中性元, 即 1_n; 再者, 前面的公式表明在 $\mathrm{M}_n(\mathrm{K})$ 内乘法关于加法是分配的.

我们称 $\mathrm{M}_n(\mathrm{K})$ 是**元素在 K 内的 n 阶矩阵环**.

即使 K 是交换的, 如果 $n \geqslant 2$(*), 环 $\mathrm{M}_n(\mathrm{K})$ 一般不是交换的. 对于 $n = 2$, 只需观察

$$\begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x & xy \\ 0 & 1 \end{pmatrix},$$

而

$$\begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}.$$

如果 $\mathrm{M}_2(\mathrm{K})$ 是交换的, 将对于任意 x, y 有 $xy = y$, 如果 K 至少具有两个元素, 这种机会是少有的.

历史上, 自然是环 $\mathrm{M}_n(\mathrm{K})$ 提供了第一批非交换环的例子.

像在所有的环内那样, 称 n 阶方阵 X 和 Y 是**交换的**, 如果

$$\mathrm{XY} = \mathrm{YX}.$$

例 4 取 X 是**对角矩阵**, 即形如

$$\mathrm{X} = \begin{pmatrix} x_1 & 0 & 0 & \cdots & 0 \\ 0 & x_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & x_n \end{pmatrix}$$

的矩阵. 我们要探索在什么条件下, 一个矩阵 $\mathrm{Y} = (y_{ij})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant n}$ 与 X 是交换的. 简单的计算指出

$$\mathrm{XY} = (x_i y_{ij}), \quad \mathrm{YX} = (y_{ij} x_j).$$

(*) 除非 $\mathrm{K} = \{0\}$.

要使两个乘积相等, 必要且充分的条件是

$$x_i y_{ij} = y_{ij} x_j \quad \text{对于 } 1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant n.$$

假定 K 是一个交换整环, 并且元素 x_i 两两不等, 考虑到 K 的交换性, 上面的条件可以写成

$$(x_i - x_j) y_{ij} = 0,$$

由于 K 是整环, 并且对于 $i \neq j$ 有 $x_i \neq x_j$, 故对于 $i \neq j$ 有 $y_{ij} = 0$, 换句话说, Y 应当是对角矩阵. 此外这个条件是充分的, 这是由于上面的计算表明元素在一个交换环内的两个对角矩阵总是交换的.

4. 同态的矩阵表示

设 L 和 M 是两个有限生成的自由的右 K-模, 而 f 是从 L 到 M 内的一个同态. 取 L 的一个基 $(a_j)_{1\leqslant j\leqslant q}$ 和 M 的一个基 $(b_i)_{1\leqslant i\leqslant p}$, 并且设

$$\mathrm{A} = (\alpha_{ij})_{1\leqslant i\leqslant p, 1\leqslant j\leqslant q}$$

是 f 关于所选定的基的矩阵. 给定一个向量

$$x = a_1\xi_1 + \cdots + a_q\xi_q,$$

f 在 x 的值

$$f(x) = b_1\eta_1 + \cdots + b_p\eta_p$$

由下式给定

$$\eta_1 = \alpha_{11}\xi_1 + \cdots + \alpha_{1q}\xi_q,$$
$$\cdots\cdots\cdots\cdots$$
$$\eta_p = \alpha_{p1}\xi_1 + \cdots + \alpha_{pq}\xi_q.$$

而这些公式显然表示以下关系成立

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1q} \\ \vdots & & \vdots \\ \alpha_{p1} & \cdots & \alpha_{pq} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_q \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_p \end{pmatrix}.$$

换句话说, 如果把每个 $x \in \mathrm{L}$ 等同于由它关于 L 的基 (a_i) 的坐标组成的列矩阵, 每个 $y \in \mathrm{M}$ 等同于由它关于 M 的基 (b_j) 的坐标组成的列矩阵, 那么向量 x 和 y 之间的关系

$$y = f(x)$$

等价于列矩阵 x 和 y 之间的关系

$$y = \mathrm{A}x.$$

这个结果使得可以以差不多机械的方式计算有限生成自由模的同态.

§§12, 13, 14 习题

1. 考虑 (元素在 $\mathbf{C}$ 内的) 矩阵

$$\mathrm{I}_1=\begin{pmatrix}1&0\\0&-1\end{pmatrix},\quad \mathrm{I}_2=\begin{pmatrix}i&0\\0&-i\end{pmatrix},\quad \mathrm{I}_3=\begin{pmatrix}0&1\\0&0\end{pmatrix},$$

$$\mathrm{I}_4=\begin{pmatrix}0&i\\0&0\end{pmatrix},\quad \mathrm{I}_5=\begin{pmatrix}0&0\\1&0\end{pmatrix},\quad \mathrm{I}_6=\begin{pmatrix}0&0\\i&0\end{pmatrix}.$$

验证下列 15 个公式:

$$\begin{array}{lll}
[\mathrm{I}_1,\mathrm{I}_3]=2\mathrm{I}_3, & [\mathrm{I}_1,\mathrm{I}_4]=2\mathrm{I}_4, & [\mathrm{I}_2,\mathrm{I}_3]=2\mathrm{I}_4,\\
[\mathrm{I}_1,\mathrm{I}_5]=-2\mathrm{I}_5, & [\mathrm{I}_1,\mathrm{I}_6]=-2\mathrm{I}_6, & [\mathrm{I}_2,\mathrm{I}_5]=-2\mathrm{I}_6,\\
[\mathrm{I}_3,\mathrm{I}_5]=\mathrm{I}_1, & [\mathrm{I}_3,\mathrm{I}_6]=\mathrm{I}_4, & [\mathrm{I}_4,\mathrm{I}_5]=\mathrm{I}_2,\\
[\mathrm{I}_2,\mathrm{I}_4]=-2\mathrm{I}_3, & [\mathrm{I}_2,\mathrm{I}_6]=2\mathrm{I}_5, & [\mathrm{I}_4,\mathrm{I}_6]=-\mathrm{I}_1,\\
[\mathrm{I}_1,\mathrm{I}_2]=0, & [\mathrm{I}_3,\mathrm{I}_4]=0, & [\mathrm{I}_5,\mathrm{I}_6]=0,
\end{array}$$

其中按照一般方式令

$$[\mathrm{X},\mathrm{Y}]=\mathrm{XY}-\mathrm{YX}.$$

找出与 6 个矩阵 $\mathrm{I}_1,\cdots,\mathrm{I}_6$ 可交换的所有二阶矩阵.

2. 对于下列矩阵验证上题的公式:

$$\mathrm{I}_1=2\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix},\quad \mathrm{I}_2=2\begin{pmatrix}0&0&0&0\\-1&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix},\quad \mathrm{I}_3=\begin{pmatrix}0&0&-1&1\\0&0&0&0\\1&0&0&0\\1&0&0&0\end{pmatrix},$$

$$\mathrm{I}_4=\begin{pmatrix}0&0&0&0\\0&0&1&-1\\0&-1&0&0\\0&-1&0&0\end{pmatrix},\quad \mathrm{I}_5=\begin{pmatrix}0&0&1&1\\0&0&0&0\\-1&0&0&0\\1&0&0&0\end{pmatrix},\quad \mathrm{I}_6=\begin{pmatrix}0&0&0&0\\0&0&1&1\\0&-1&0&0\\0&1&0&0\end{pmatrix}.$$

3. 求与矩阵

$$\begin{pmatrix}1&0&0\\0&1&0\\3&1&2\end{pmatrix}$$

可交换的所有三阶矩阵 (取基础环或为 $\mathbf{C}$, 或为任意交换域, 或为一个任意环).

4. 设 K 是一个交换环. 证明把每个矩阵 $\begin{pmatrix}a&b\\c&d\end{pmatrix}$ 变换为矩阵

$$\begin{pmatrix}a&0&b&0\\0&a&0&b\\c&0&d&0\\0&c&0&d\end{pmatrix}$$

的从 $M_2(K)$ 到 $M_4(K)$ 的映射是环 $M_2(K)$ 到 $M_4(K)$ 的一个子环上的同构. 用模的术语解释这个结果.

5. 计算以下三个矩阵的乘积:

$$\begin{pmatrix} 0 & 2 & -1 \\ -2 & -1 & 2 \\ 3 & -2 & -1 \end{pmatrix}, \quad \begin{pmatrix} 70 & 34 & -107 \\ 52 & 26 & -68 \\ 101 & 50 & -140 \end{pmatrix}, \quad \begin{pmatrix} 27 & -18 & 10 \\ -46 & 31 & -17 \\ 3 & 2 & 1 \end{pmatrix}.$$

取基础环为模 7 整数的环 $\mathbf{Z}/7\mathbf{Z}$ 进行同样的计算.

6. 计算以下 n 阶方阵的立方

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

7. 设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 是其元素在一个任意交换环内的二阶矩阵. 证明

$$A^2 - (a+d)A + (ad - bc)1_2 = 0.$$

8. 给定系数在一个交换环 K 内的一个方阵

$$A = (\alpha_{ij})_{1 \leqslant i,j \leqslant n},$$

称标量

$$\mathrm{Tr}(A) = \alpha_{11} + \alpha_{22} + \cdots + \alpha_{nn},$$

即 A 的对角线元素之和为 A 的**迹**. 证明对于任意 n 行 n 列矩阵 A 和 B 有

$$\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B), \quad \mathrm{Tr}(AB) = \mathrm{Tr}(BA).$$

假定 $K = \mathbf{C}$. 由此推出不可能找到 n 阶方阵 X 和 Y, 使得

$$XY - YX = 1_n.$$

9. 设 K 是一个交换环, 而 d 是 K 的一个元素. 证明矩阵

$$\begin{pmatrix} x & dy \\ y & x \end{pmatrix}$$

(其中的 x 和 y 是 K 的任意元素) 组成 $M_2(K)$ 的一个子环 L, 并且 L 同构于 §9 的环 $K[\sqrt{d}]$. 当 $K = \mathbf{R}, d = -1$ 时情形如何?

10. 称一个元素在环 K 中的 n 阶方阵 X 是**幂零的**, 如果存在一个整数 $r \geqslant 1$, 使得 $X^r = 0$; 称 X 是**幂幺的**, 如果 $1_n - X$ 是幂零的. 以下假定 $K = \mathbf{C}$. 给定一个幂零方阵 N 和一个幂幺方阵 U, 令 (§8, 习题 2)

$$\exp(N) = 1 + \frac{N}{1!} + \frac{N^2}{2!} + \cdots + \frac{N^k}{k!} + \cdots,$$
$$\log(U) = -\frac{1-U}{1} - \frac{(1-U)^2}{2} - \cdots - \frac{(1-U)^k}{k} - \cdots,$$

取

$$\mathrm{N}=\begin{pmatrix}0 & a & c\\ 0 & 0 & b\\ 0 & 0 & 0\end{pmatrix},\quad \mathrm{U}=\begin{pmatrix}1 & x & z\\ 0 & 1 & y\\ 0 & 0 & 1\end{pmatrix},$$

验证 N 是幂零的, U 是幂幺的, 并且有关系

$$\exp(\log(\mathrm{U}))=\mathrm{U},\quad \log(\exp(\mathrm{N}))=\mathrm{N}.$$

[不利用 §8 习题 2 的一般结果, 实际计算矩阵 $\exp(\log(\mathrm{U}))$ 和 $\log(\exp(\mathrm{N}))$.]

11. 对于所有的复数 t, 令

$$\mathrm{U}(t)=\begin{pmatrix}1 & t & 2t+2t^2 & 3t+\frac{17}{2}t^2+4t^3\\ 0 & 1 & 4t & 5t+12t^3\\ 0 & 0 & 1 & 6t\\ 0 & 0 & 0 & 1\end{pmatrix},$$

证明对于任意 $s,t\in\mathbf{C}$, 有

$$\mathrm{U}(s)\mathrm{U}(t)=\mathrm{U}(s+t),$$

和 $\mathrm{U}(t)=\exp(t\mathrm{N})$, 其中 N 是一个要求算出的幂零矩阵.

¶12. 设 z 是一个代数数, 即 (§11, 习题 11) 一个带不全为零的有理系数的代数方程的一个根.

a) 证明存在一个整数 $n\geqslant 1$ 和有理数 $a_0,\cdots,a_{n-1}$, 使得

$$z^n=a_0+a_1z+\cdots+a_{n-1}z^{n-1}.$$

b) 设 K 是由 $\mathbf{Q}$ 和 z 生成的 $\mathbf{C}$ 的一个子环, 证明看作 $\mathbf{Q}$ 上的向量空间, K 是由元素 $1,z,\cdots,z^{n-1}$ 生成的.

c) 假定 n 是前面的 n 中的最小者. 证明 $1,z,\cdots,z^{n-1}$ 组成看作 K 作为 $\mathbf{Q}$ 上的向量空间的一个基.

d) 设 f 是由

$$f(u)=zu\quad 对于所有\ u\in\mathrm{K}$$

定义的从 K 到 K 内的映射. 证明这是看作 $\mathbf{Q}$ 上的向量空间的 K 的自同态. 计算 f 关于问题 c) 中所定义的基的矩阵.

¶13. 设 L 和 M 是环 K 上的两个左模, 假定给定 L 的一个子模 L′ 和从 L 到 M 内的一个同态 f. 证明下列两个条件是等价的:

a) L′ 包含于 f 的核内, 即对于所有的 $x\in\mathrm{L}'$ 有 $f(x)=0$;

b) f 是从 L 到商模 L/L′ (§10, 习题 10) 上的典范映射和从 L/L′ 到 M 内的一个同态的复合.

14. 设 K 是一个环, L, M 和 N 是三个左 K-模, f 是从 L 到 M 内的一个同态, 而 p 从 L 到 N 内的一个同态. 假定 p 是满射. 证明下列两个条件是等价的:

a) $\mathrm{Ker}(f)\supset\mathrm{Ker}(p)$;

b) f 是 p 和从 N 到 M 内的一个同态的复合.

15. 设 L 是通常空间内的起点为 O 的所有向量组成的实向量空间, 而 L' 是由位于过 O 的一条直线 (对应的, 一张平面) 上的所有向量组成的向量子空间. 对于所有的 $x \in L$, 用 $f(x)$ 表示 x 在 L' 上的投影向量. 证明从 L 到 L' 内的映射 f 是线性的. 它的核是什么?

¶¶16. 设 K 是一个环. 称一个左 K-模 M 是**单的**, 或**不可约的**, 如果它不缩减为零, 并且 M 的仅有的子模是 $\{0\}$ 和 M 本身.

a) 假定 M 非零, M 是单的, 必须并且只需对于任意 $a, b \in M$, 其中的 $a \neq 0$, 存在一个 $\lambda \in K$, 使得 $b = \lambda a$. 如果 K 是一个域, 由此推出所有的单的 K-模同构于 K.

b) 设 M 是一个单的 K-模. 在 M 内取一个元素 $a \neq 0$, 用 I 表示使得 $\lambda a = 0$ 的 $\lambda \in K$ 的集合. 证明 I 是 K 的一个*极大的* (§8, 习题 7) 左理想. 证明从 K 到 M 内的映射 $\lambda \to \lambda a$ 是从 K 到 K/I 的典范映射和从左 K-模 K/I 到 M 上的一个同构的复合.

证明反过来对于 K 的所有极大左理想 I, 左 K-模 K/I 是单的.

c) 设 L 和 M 是单的左 K-模, 而 f 是从 L 到 M 内的一个同态. 证明, 如果 f 不是零, 那么这是从 L 到 M 上的一个同构 (Schur *引理*). [考察 f 的核与像.]

d) 设 L 是一个单的左 K-模. 证明 L 的自同态环是一个域 (一般是非可交换的).

¶17. 设 $u = (a, b)$ 是 $\mathbf{Z}$-模 $\mathbf{Z}^2$ 的一个元素.

a) 证明如果存在 $\mathbf{Z}^2$ 的含有 u 的一个基, 那么在所考虑的模上存在一个线性型 f, 使得 $f(u) = 1$. 由此推出整数 a 和 b 是互素的.

b) 反之假定 a 和 b 是互素的. 证明存在 $\mathbf{Z}^2$ 上的一个线性型 f, 使得 $f(u) = 1$. 证明存在一个向量 v, 使得 $\mathrm{Ker}(f)$ 是 v 的整数倍的集合. 证明 u 和 v 组成 $\mathbf{Z}^2$ 的一个基.

c) 取 $u = (6, 35)$. 求一个向量 v, 使得 u 和 v 组成 $\mathbf{Z}^2$ 的一个基.

18. 设 L 和 M 是任意环 K 上的两个左模, 而

$$f : M \to L$$

是一个*满射的*同态. 假定 L 是有限生成自由的. 证明存在一个同态

$$g : L \to M,$$

使得

$$f \circ g = id.$$

¶19. 设 L 和 M 是一个环 K 上的两个左模, L' 和 M' 分别是 L 和 M 的子模, 而 u 是从 L 到 M 内的一个同态, 使得 $u(L') \subset M'$. 用 p 和 q 分别表示从 L 到 L/L' 上的和从 M 到 M/M' 上的典范映射 (§§10, 11, 习题 10). 证明存在一个同态

$$\bar{u} : L/L' \to M/M',$$

使得有

$$q \circ u = \bar{u} \circ p$$

(说 $\bar{u}$ 是由 u 被过渡到商诱导的同态). 在什么条件下 $\bar{u}$ 是一个单射, 满射或双射?

§15 逆矩阵和基的变换

1. 模的自同构群

在 §12 第 1 小节曾经提到, 称所有从一个模 M 到 M 上双射的同态, 即从 M 到 M 上的同构, 为模 M 的自同构. M 的自同构是集合 M 的置换的特殊情形. §12 的定理 1 表明, 如果 u 和 v 是 M 的自同构, 则映射 $u \circ v^{-1}$ 也是; 因此, 模 M 的自同构的集合

$$\mathrm{GL(M)}$$

是集合 M 的置换群 $\mathfrak{S}(\mathrm{M})$ 的一个子群. 我们说 GL(M) 是**模 M 的自同构群**, 或 **M 的一般线性群**. 形如 GL(M) 的群和它们的子群在群的一般理论的发展中发挥着一流的作用.

我们注意到还可以从模 M 的自同态环 Hom(M, M) 出发定义群 GL(M). 显然有

$$\mathrm{GL(M)} \subset \mathrm{Hom(M, M)},$$

而 GL(M) 的元素正是环 Hom(M, M) 的*可逆元*. 事实上, 如果一个自同态 u 是可逆的, 则存在一个自同态 v, 使得 $u \circ v = v \circ u = j_{\mathrm{M}}$, 因此 u 是双射的, 从而属于 GL(M); 其逆是显然的. 总之, GL(M) 正是 §8 中注 1 意义下的环 Hom(M, M) 的*可逆元群*.

2. 群 GL(n, K)

一个矩阵 $\mathrm{U} \in \mathrm{M}_n(\mathrm{K})$ 是**可逆的**, 如果存在一个矩阵 $\mathrm{V} \in \mathrm{M}_n(\mathrm{K})$, 使得

$$\mathrm{UV} = \mathrm{VU} = 1_n.$$

换句话说, 如果 U 是环 $\mathrm{M}_n(\mathrm{K})$ 的可逆元. 此时矩阵 V 是唯一的, 记为 U^{-1}, 并且称为 U 的**逆矩阵**. 把元素在 K 内的 n 阶可逆方阵的集合记作

$$\mathrm{GL}(n, \mathrm{K}),$$

配备了运算 $(\mathrm{U}, \mathrm{V}) \to \mathrm{UV}$, 这个集合是一个群, 称为**环 K 上的 n 个变量的一般线性群**. 这正是环 $\mathrm{M}_n(\mathrm{K})$ 的可逆元素的乘法群.

设 M 是一个*有限生成的自由模*, 取 M 的一个基 $a_1, \cdots, a_n$, M 的每个自同态 f 捆绑它关于这个基的一个矩阵 (§12, 第 3 小节), 把它记为 $\mathrm{A}(f)$, 我们得到从环 Hom(M, M) 到元素在 K 内的 n 阶方阵环 $\mathrm{M}_n(\mathrm{K})$ 上的一个*双射*

$$f \to \mathrm{A}(f),$$

而定义矩阵的和与乘积的方式表明有关系

$$\mathrm{A}(f+g) = \mathrm{A}(f) + \mathrm{A}(g), \quad \mathrm{A}(fg) = \mathrm{A}(f)\mathrm{A}(g), \quad \mathrm{A}(j_{\mathrm{M}}) = 1_n,$$

换句话说, 映射 $f \to \mathrm{A}(f)$ 是从环 $\mathrm{Hom}(\mathrm{M},\mathrm{M})$ 到环 $\mathrm{M}_n(\mathrm{K})$ 上的一个同构.

由于一个环 U 到一个环 V 上的同构显然映射 U 的可逆元的集合 U^* 到 V 的可逆元的集合 V^*, 我们得到结论: M 的一个自同态 f 是 M 的一个同构, 必须并且只需它的矩阵 $\mathrm{A}(f)$ 是环 $\mathrm{M}_n(\mathrm{K})$ 的可逆元. 换句话说, 关系

$$f \in \mathrm{GL}(\mathrm{M}) \quad 和 \quad \mathrm{A}(f) \in \mathrm{GL}(n,\mathrm{K})$$

是等价的.

如果像在 §14 中的第 3 小节那样, 把 $\mathrm{M}_n(\mathrm{K})$ 与右 K-模 K^n 的自同态环等同, 那么我们发现 $\mathrm{GL}(n,\mathrm{K})$ 就跟 K^n 的自同构群等同, 即跟从 K^n 到 K^n 内所有的映射

$$(\xi_1,\cdots,\xi_n) \to (\eta_1,\cdots,\eta_n)$$

的群同构, 这些映射是双射的并且是线性的, 即它由形如

$$\eta_i = \alpha_{i1}\xi_1 + \cdots + \alpha_{in}\xi_n \quad (1 \leqslant i \leqslant n)$$

的公式给定

3. 例子: 群 $\mathrm{GL}(1,\mathrm{K})$ 和 $\mathrm{GL}(2,\mathrm{K})$

对于 $n=1$, 环 $\mathrm{M}_n(\mathrm{K})$ 等同于环 K 本身, 因此群 $\mathrm{GL}(1,\mathrm{K})$ 缩减为 K 的可逆元的乘法群 K^*.

为了研究群 $\mathrm{GL}(2,\mathrm{K})$, 我们假定 K 是交换的. 为了一个矩阵

$$\mathrm{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

具有一个逆矩阵

$$\mathrm{A}^{-1} = \begin{pmatrix} x & z \\ y & t \end{pmatrix},$$

必须存在 $x,y,z,t \in \mathrm{K}$, 使得

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & z \\ y & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{1}$$

即

$$\begin{cases} ax+by=1, \\ cx+dy=0, \end{cases} \qquad \begin{cases} az+bt=0, \\ cz+dt=1. \end{cases} \tag{2}$$

为了找到方程 (2) 有解的必要并且充分的条件, 我们引进元素在一个交换环 K 内的二阶方阵的**行列式**概念: 对于一个矩阵

$$\mathrm{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

称数量 $ad-bc$ 为它的行列式, 记作

$$\det(\mathrm{A}) \quad \text{或} \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

后面我们将推广这个定义到任意 n 阶方阵.

当前, 我们知道对于任意 $\mathrm{A}, \mathrm{B} \in \mathrm{M}_2(\mathrm{K})$ 有

$$\det(\mathrm{AB}) = \det(\mathrm{A})\det(\mathrm{B})$$

就足够了. 这个关系事实上等价于等式

$$(ad-bc)(xt-yz) = (ax+by)(cz+dt) - (az+bt)(cx+dy),$$

考虑到 K 的交换性, 读者不难验证此式.

做好了这些准备, 由于单位矩阵 1_2 的行列式显然是 1, 关系

$$\det(\mathrm{A})\det(\mathrm{A}^{-1}) = 1$$

表明为了矩阵 A 是可逆的, 必须它的行列式是 K 的可逆元. 反之, 如果 $ad-bc$ 在 K 内是可逆的, 考虑 (2) 里的 x 和 y 的关系, 容易验证如果取

$$x = (ad-bc)^{-1}d, \quad y = -(ad-bc)^{-1}c,$$

它们将满足, 而在 (2) 里的 z 和 t 的关系对于

$$z = -(ad-bc)^{-1}b, \quad t = (ad-bc)^{-1}a$$

是满足的. 容易验证这样构造的矩阵

$$\begin{pmatrix} x & z \\ y & t \end{pmatrix}$$

是 A 的右逆矩阵和左逆矩阵.

总之, 如果 K 是一个交换环, 当且仅当 $ad-bc$ 在 K 内是可逆的, 矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 是可逆的.

举例来说, 如果 K 是一个交换域, 那么 GL(2, K) 由满足条件

$$ad-bc \neq 0$$

的矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 组成. 反之, 群 $\mathrm{GL}(2, \mathbf{Z})$ 由元素为整数并且满足条件

$$ad-bc = 1 \quad \text{或} \quad ad-bc = -1$$

的矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 组成.



注 1 设 L 是一个环, 而 K 是 L 的一个子环. 显然可以把 $\mathrm{M}_n(\mathrm{K})$ 看作 $\mathrm{M}_n(\mathrm{L})$ 的一个子环. 那么有可能一个矩阵

$$\mathrm{U} \in \mathrm{M}_n(\mathrm{K})$$

在环 $\mathrm{M}_n(\mathrm{L})$ 内是可逆的, 在环 $\mathrm{M}_n(\mathrm{K})$ 内却不然: 取仅有一个元素的矩阵 (2) 或矩阵 $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, 看作元素在域 $\mathbf{Q}$ 内的矩阵, 它们是可逆的, 但是看作元素在环 $\mathbf{Z}$ 内的矩阵, 它们不是可逆的. 如果不明确指定基础环, 那么可逆矩阵的概念就有引起混淆的危险.

不过当所考虑的基础环是域时这类困难不会出现. 即如果 L 是一个域, 而 K 是其子域; 那么如果矩阵 $\mathrm{U} \in \mathrm{M}_n(\mathrm{K})$ 在环内 $\mathrm{M}_n(\mathrm{L})$ 是可逆的, 那么它在 $\mathrm{M}_n(\mathrm{K})$ 内也是可逆的. 在 §20 习题 20 内有证明. 如果 L 是交换的, 还可以利用 §23 的定理 8 的推论 1.

4. 基的变换: 过渡矩阵

设 M 是一个有限生成自由右 K-模, 考虑 M 的有 n 个元素的两个基 $(a_1, \cdots, a_n)$ 和 $(b_1, \cdots, b_n)$ [如果 K 是一个域 (§19, 定理 6), 或是交换域 (§23, 定理 5 的推论), 这个条件总是满足的]. 我们打算对于每个 $x \in \mathrm{M}$, 通过它关于第一个基的坐标计算它关于第二个基的坐标.

为此, 考虑由

$$u(e_i) = a_i \quad v(e_i) = b_i \quad (1 \leqslant i \leqslant n) \tag{3}$$

定义的同态

$$u, v : \mathrm{K}^n \to \mathrm{M},$$

其中 $e_1, \cdots, e_n$ 是 K^n 的典范基. (a_i) 和 (b_i) 是 M 的基, 必须且只需 u 和 v 同构 (§12, 定理 3 的推论 1). 用 $\xi_1, \cdots, \xi_n$ 表示 x 关于 (a_i) 的坐标, 用 $\eta_1, \cdots, \eta_n$ 表示 x 关于 (b_i) 的坐标, 我们有关系

$$x = u(\xi_1, \cdots, \xi_n) = v(\eta_1, \cdots, \eta_n).$$

引进 K^n 的同构 (§12, 定理 1)

$$w = v^{-1} \circ u, \tag{4}$$

则有

$$(\eta_1, \cdots, \eta_n) = w(\xi_1, \cdots, \xi_n);$$

用

$$(\alpha_{ij})_{1\leqslant i\leqslant n,1\leqslant j\leqslant n}$$

表示 w 关于 K^n 的典范基的矩阵, 则得公式

$$\eta_i = \alpha_{i1}\xi_1 + \cdots + \alpha_{in}\xi_n \quad (1 \leqslant i \leqslant n), \tag{5}$$

此式就解决了开头提出的问题. 称为**坐标变换公式**, 而出现在公式中的矩阵 (α_{ij}) 称为**从基** (a_i) **到基** (b_i) **的过渡矩阵**. 由于这是 K^n 的一个自同构, 故

$$(\alpha_{ij}) \in \mathrm{GL}(n, \mathrm{K}).$$

我们观察到关系 (4) 还可以写成

$$v = u \circ w^{-1}. \tag{6}$$

自同构 w^{-1} (关于 K^n 的典范基) 的矩阵必然是过渡矩阵的逆矩阵

$$(\beta_{ij}) = (\alpha_{ij})^{-1}.$$

我们有关系

$$w^{-1}(e_j) = e_1\beta_{1j} + \cdots + e_n\beta_{nj},$$

考虑到 (3) 即得

$$\begin{aligned} b_j = v(e_j) &= u[w^{-1}(e_j)] \\ &= u(e_1\beta_{1j} + \cdots + e_n\beta_{nj}) \\ &= u(e_1)\beta_{1j} + \cdots + u(e_n)\beta_{nj}, \end{aligned}$$

即

$$b_j = a_1\beta_{1j} + \cdots + a_n\beta_{nj} \quad (1 \leqslant j \leqslant n). \tag{7}$$

细心观察出现在关系 (5) 和 (7) 中的 (α_{ij}) 和 (β_{ij}), 它们是不同的, 而是*互为逆矩阵*.

类似的计算证明了关系

$$a_j = b_1\alpha_{1j} + \cdots + b_n\alpha_{nj}. \tag{7'}$$

注 2 最简单的情形是 $n = 1$. 此时有仅含有一个向量的两个基 (a) 和 (b), 并且有关系

$$x = a\xi = b\eta;$$

令

$$\eta = \alpha\xi,$$

则对于任意 ξ 应当有 $a\xi = b\alpha\xi$, 于是有 $a = b\alpha$, 即

$$b = a\alpha^{-1}.$$

换言之, 如果基向量 a 换为 $a\alpha$, x 的坐标 ξ 就换为 $\alpha^{-1}\xi$, 这是符合常理的, 因为乘积 $a\xi$ 应当保持常值.

前面的结果使得有可能从 M 的一个基出发构造所有其他的基.

定理 1　设 $a_1, \cdots, a_n$ 是右 K-模 M 的一个基. 向量

$$b_j = a_1\beta_{1j} + \cdots + a_n\beta_{nj} \quad (1 \leqslant j \leqslant n)$$

组成 M 的一个基, 必须且只需矩阵 $(\beta_{ij})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant n}$ 在 $\mathrm{M}_n(\mathrm{K})$ 内是可逆的.

事实上, 考虑由公式 (3) 给定的同态 $u, v : \mathrm{K}^n \to \mathrm{M}$. 我们知道, 由于 (a_i) 组成一个基, u 是双射的. 为了说明 (b_i) 组成一个基, 应当说明 v 是双射的, 而这显然归结为说明 $u^{-1}\circ v$ 是 K^n 的自同构. 而我们有

$$\begin{aligned} u^{-1}\circ v(e_j) &= u^{-1}(b_j) = u^{-1}(a_1\beta_{1j} + \cdots + a_n\beta_{nj}) \\ &= u^{-1}(a_1)\beta_{1j} + \cdots + u^{-1}(a_n)\beta_{nj} \\ &= e_1\beta_{1j} + \cdots + e_n\beta_{nj}, \end{aligned}$$

因此 (β_{ij}) 恰好是 K^n 的自同态 $u^{-1}\circ v$ 的矩阵. 由于当且仅当 K^n 的一个自同态的矩阵在 $\mathrm{GL}(n, \mathrm{K})$ 内, 它是双射的, 因此定理证毕.

定理 1 的条件还可以这样得到: 引进 M 的自同态 f:

$$f(a_i) = b_i \quad (1 \leqslant i \leqslant n)$$

(它的存在性从 §12 定理 3 得到). 诸 b_i 组成 M 的一个基, 必须且只需 f 是 M 的自同构; 而矩阵 (β_{ij}) 正是 f 关于 M 的基 (a_i) 的矩阵.

这个结果是极其显然的, 重新引进上面的同态 u 和 v, 我们有

$$f = v \circ u^{-1},$$

由于 u 是双射, 那么 v 是双射归结为 f 是双射.

5. 基的变换对于一个同态的矩阵的影响

设 L 和 M 是两个有限生成自由右 K-模, 而 $f : \mathrm{L} \to \mathrm{M}$ 是一个同态. 又设

$$(a'_1, \cdots, a'_q) \quad 和 \quad (a''_1, \cdots, a''_q)$$

是 L 的含有 q 个元素的两个基, 而

$$(b'_1, \cdots, b'_p) \quad 和 \quad (b''_1, \cdots, b''_p)$$

是 M 的含有 p 个元素的两个基. 再设 A' 是 f 关于基 (a'_j) 和 (b'_i) 的矩阵, 而 A'' 是 f 关于基 (a''_j) 和 (b''_i) 的矩阵. 我们打算通过 A', 基 (a'_j) 到基 (a''_j) 的过渡矩阵

$$\mathrm{U} \in \mathrm{GL}(q, \mathrm{K}),$$

和基 (b'_i) 到基 (b''_i) 的过渡矩阵

$$\mathrm{V} \in \mathrm{GL}(p, \mathrm{K})$$

计算 A''.

为此考虑同态 $u', u'' : \mathrm{K}^q \to \mathrm{L}$, 它们分别映射 K^q 的典范基到 L 的基 (a'_j) 和 (a''_j), 同态 $v', v'' : \mathrm{K}^p \to \mathrm{M}$, 它们分别映射 K^p 的典范基到 M 的基 (b'_i) 和 (b''_i). 我们有关系 $u' = u'' \circ u, v' = v'' \circ v$, 这里 u (对应的, v) 是 K^q (对应的, K^p) 的自同构, 根据前一小节, 它关于 K^q (对应的, K^p) 的典范基的矩阵恰好是 U (对应的, V). 此外, 如果引进由

$$f' = v'^{-1} \circ f \circ u', \quad f'' = v''^{-1} \circ f \circ u''$$

定义的同态

$$f', f'' : \mathrm{K}^q \to \mathrm{K}^p,$$

那么根据 §12 的注 2, 上面定义的矩阵 A' 和 A'' 正是 f' 和 f'' 关于 K^q 和 K^p 的典范基的矩阵. 引进上面定义的 u 和 v 则有

$$\begin{aligned} f' &= (v'' \circ v)^{-1} \circ f \circ (u'' \circ u) \\ &= v^{-1} \circ v''^{-1} \circ f \circ u'' \circ u = v^{-1} \circ f'' \circ u. \end{aligned}$$

由于两个同态的复合翻译成它们的矩阵的乘积, 取 u, v, f', f'' 关于典范基的矩阵即得关系

$$\mathrm{A}' = \mathrm{V}^{-1}\mathrm{A}''\mathrm{U},$$

此式解决了本小节开头所提的问题:

定理 2 设 L 和 M 是两个有限生成自由右 K-模, 而 $f : \mathrm{L} \to \mathrm{M}$ 是一个同态. 又设 A' 是 f 关于 L 的基 $(a'_j)_{1 \leqslant j \leqslant q}$ 和 M 的基 $(b'_i)_{1 \leqslant i \leqslant p}$ 的矩阵, 而 A'' 是 f 关于 L 的基 $(a''_j)_{1 \leqslant j \leqslant q}$ 和 M 的基 $(b''_i)_{1 \leqslant i \leqslant p}$ 的矩阵. 最后设 U 是从基 (a'_j) 到基 (a''_j) 的过渡矩阵, 而 V 是从 (b'_i) 到 (b''_i) 的过渡矩阵. 则有关系

$$\mathrm{A}' = \mathrm{V}^{-1}\mathrm{A}''\mathrm{U}.$$

当 $\mathrm{L} = \mathrm{M}$ 时, 在前面可以假定基 (b'_i) 和 (a'_j) 相同, 而基 (b''_i) 和 (a''_j) 相同, 这时显然 $\mathrm{U} = \mathrm{V}$; 故得

推论　设 L 是一个有限生成自由右 K-模, 而 f 是 L 的一个自同态, 又设

$$(a'_i)_{1\leqslant i\leqslant p} \quad 和 \quad (a''_i)_{1\leqslant i\leqslant p}$$

是 L 的有同样个数元素的两个基. A′ 是 f 关于 L 基 (a'_i) 的矩阵, 而 A″ 是 f 关于基 (a''_i) 的矩阵. 则有

$$A'' = UA'U^{-1},$$

其中 U 是从基 (a'_i) 到基 (a''_i) 的过渡矩阵.

这一结果导出一个重要概念: 给定矩阵

$$A', A'' \in M_p(K),$$

如果存在一个矩阵

$$U \in GL(p, K),$$

使得

$$A'' = UA'U^{-1},$$

则称 A′ 和 A″ 是**相似的**.



注 3　总存在 (定理 1) 基的一个变换, 使得一个任意选择的可逆矩阵为其过渡矩阵. 于是发现定理 2 具有一个逆定理: 如果预先给定同态 f 和基 (a'_j), (b'_i), 从而给定了矩阵 A′, 那么对于任意 $U \in GL(q, K)$ 和 $V \in GL(p, K)$, 在 L 和 M 里存在基, 使得 f 关于它们的矩阵是

$$V^{-1}A'U.$$

对于推论有一个类似的结果.

注 4　我们刚给的定理 2 的证明, 避免了所有明晰的计算, 但是不得不以 “原型” 模 K^p 和 K^q 为中介来过渡, 这可能使初学的读者为难.

还可以如下证明定理 2: 设 U 和 V 是过渡矩阵, 令

$$U^{-1} = (\omega_{kl})_{1\leqslant k,l\leqslant q}, \quad V^{-1} = (\rho_{ij})_{1\leqslant i,j\leqslant p};$$

根据第 4 小节我们有

$$a''_i = \sum_k a'_k\omega_{kl}, \quad b''_j = \sum_i b'_i\rho_{ij}; \tag{8}$$

再令

$$A' = (\alpha'_{ik})_{1\leqslant i\leqslant p, 1\leqslant k\leqslant q},$$
$$A'' = (\alpha''_{jl})_{1\leqslant j\leqslant p, 1\leqslant l\leqslant q},$$

则有

$$f(a'_k) = \sum_i b'_i \alpha'_{ik}, \quad f(a''_l) = \sum_j b''_j \alpha''_{jl}. \tag{9}$$

考虑到 (8), (9) 的第二个关系改写为

$$\sum_k f(a'_k)\omega_{kl} = \sum_j \sum_i b'_i \rho_{ij} \alpha''_{jl},$$

根据 (9) 的第一个关系,

$$\sum_k \sum_i b'_i \alpha'_{ik} \omega_{kl} = \sum_j \sum_i b'_i \rho_{ij} \alpha''_{jl}.$$

由于向量 b'_i 是线性无关的, 它们在等式两端的系数应当相等, 这就对于任意 i 和 l 给出关系

$$\sum_k \alpha'_{ik} \omega_{kl} = \sum_j \rho_{ij} \alpha''_{jl}.$$

而左端是矩阵 $\mathrm{A}'\mathrm{U}^{-1}$ 的指标为 i, l 的元素, 右端是矩阵 $\mathrm{V}^{-1}\mathrm{A}''$ 指标为 i, l 的元素; 故得

$$\mathrm{A}'\mathrm{U}^{-1} = \mathrm{V}^{-1}\mathrm{A}'',$$

由此即得要找的关系

$$\mathrm{A}'' = \mathrm{VAU}^{-1}.$$

这类计算在线性代数和 "张量" 理论里是屡见不鲜的, 并且理所当然地给那些相信只有爱因斯坦本人能够理解他自己工作的人以深刻的印象. 今天, 大多数数学家更喜欢摆脱泛滥的指标, 而代之以几何的或更准确地说是概念的推理, 好处是变得更简单. 但是, 大部分物理学家仍然使用类似于本注里所陈述的方法 (这让人难以理解, 物理学家应当比数学家对于 "几何的" 或 "物理的" 对象更感兴趣, 而不是对于这些对象的坐标更感兴趣, 尤其当不是为了做实际的计算时). 所以熟悉关于指标和求和号 Σ 的计算还是有用的, 尽管它们在理论上显得多余.

§15 习题

1. 基础环是 **R**, 求下列矩阵的逆矩阵(*):

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}; \quad \begin{pmatrix} 2 & 5 & 7 \\ 6 & 3 & 4 \\ 5 & -2 & -3 \end{pmatrix}; \quad \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}; \quad \begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}.$$

(*) 更超前的读者可以利用 Cramer 公式解这些习题.

2. 计算以下矩阵的逆矩阵:

$$\begin{pmatrix} 1 & a & a^2 & \cdots & a^n \\ 0 & 1 & a & \cdots & a^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

3. 设 N 是一个其元素在一个环内的幂零方阵 (即它的一个幂是零). 证明矩阵 $1-\mathrm{N}$ 是可逆的, 并且

$$(1-\mathrm{N})^{-1} = 1 + \mathrm{N} + \mathrm{N}^2 + \cdots.$$

应用: 计算以下矩阵的逆矩阵:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

4. 计算以下矩阵的逆矩阵:

$$\begin{pmatrix} 1+a_1 & 1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1+a_n \end{pmatrix}.$$

¶5. 设 $\omega = \cos(2\pi/n) + i \cdot \sin(2\pi/n)$, 证明矩阵

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)^2} \end{pmatrix}$$

的逆矩阵由把这个矩阵里的 ω 换为 ω^{-1}, 并且把这样得到的矩阵除以 n 而得到.

6. 求满足等式的三阶方阵 X(*)

$$\begin{pmatrix} 2 & -3 & 1 \\ 4 & -5 & 2 \\ 5 & -7 & 3 \end{pmatrix} \mathrm{X} \begin{pmatrix} 9 & 7 & 6 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -2 \\ 18 & 12 & 9 \\ 23 & 15 & 11 \end{pmatrix}.$$

7. 证明向量

$$(1,2,1), \quad (2,3,3), \quad (3,7,1)$$

(*) 见前一页的脚注.

组成 $\mathbf{R}^3$ 的一个基, 同样

$$(3,1,4),\quad (5,2,3),\quad (1,1,-6)$$

也组成 $\mathbf{R}^3$ 的一个基, 并且计算从第一个基到第二个基的过渡矩阵.

对于 $\mathbf{R}^4$ 向量

$$(1,1,1,1),\quad (1,2,1,1),\quad (1,1,2,1),\quad (1,3,2,3)$$

和

$$(1,0,3,3),\quad (-2,-3,-5,-4),\quad (2,2,5,4),\quad (-2,-3,-4,-4)$$

解同样的问题.

8. 证明矩阵

$$\begin{pmatrix} x+y & 4y \\ -y & x-y \end{pmatrix}$$

组成环 $\mathrm{M}_2(\mathbf{Q})$ 的一个子环, 其中的 x 和 y 是任意有理数.

9. 矩阵

$$\begin{pmatrix} x & y & z \\ 2z & x & y \\ 2y & 2z & x \end{pmatrix}$$

组成环 $\mathrm{M}_3(\mathbf{Q})$ 的一个子环, 其中的 x,y,z 是任意有理数.

¶10. 设 K 是一个交换环, p 和 q 是 K 的给定元素. 考虑环

$$\mathrm{L} = \mathrm{K}[\sqrt{p}]$$

和形如

$$z = \begin{pmatrix} x & qy \\ \bar{y} & \bar{x} \end{pmatrix}, \text{ 其中 } x, y \in \mathrm{L}$$

的矩阵的集合 $\mathrm{M} \subset \mathrm{M}_2(\mathrm{L})$ (关于记号, 参见 §9).

a) 证明 M 是 $\mathrm{M}_2(\mathrm{L})$ 的一个子空间.

b) 对于 M 的所有矩阵

$$z = \begin{pmatrix} x & qy \\ \bar{y} & \bar{x} \end{pmatrix}, \tag{1}$$

令

$$z^* = \begin{pmatrix} \bar{x} & -qy \\ -\bar{y} & x \end{pmatrix},$$

证明对于任意 $z_1, z_2 \in \mathrm{M}$, 有 $(z_1 z_2)^* = (z_2)^*(z_1)^*$.

c) 对于矩阵 (1), 计算 z^*z, 并且证明 z 是环 M 的可逆元当且仅当

$$\mathrm{N}(z) = \bar{x}x - q\bar{y}y$$

是环 L 的可逆元.

d) 假定 K 是一个交换域, 并且 p 不是 K 内的一个平方. 证明下列断言是等价的:

(i) 环 M 是一个域;

(ii) 不存在 K 的任何一对元素 x, y, 使得

$$q = x^2 - py^2.$$

e) 假定 $\mathrm{K} = \mathbf{R}$. 证明 M 是一个域当且仅当有

$$p < 0, \quad q < 0.$$

并且证明这样得到的域同构于取 $p = q = -1$ 将得到的域 (称为**四元数域**, 历史上的第一个非交换域).

f) 假定 $\mathrm{K} = \mathbf{Q}$, 而 p 和 q 是整数. 证明, M 是一个域, 必须并且只需方程

$$px^2 + qy^2 = z^2$$

不具有 (0, 0, 0) 以外的任何整数解 (x, y, z). 证明这个条件在下列情形是满足的, 举例说:

$$\begin{aligned} &p = 5, \quad q = 2(\mathrm{mod}\ 5);\\ &p = 5, \quad q = 3(\mathrm{mod}\ 5);\\ &p = 11, \quad q = 2, 6, 7, 8, \text{ 或 } 10(\mathrm{mod}\ 11). \end{aligned}$$

¶11. 证明形如

$$\begin{pmatrix} x & -y & -z & -t \\ y & x & -t & z \\ z & t & x & -y \\ t & -z & y & x \end{pmatrix}$$

的矩阵组成 $\mathrm{M}_4(\mathbf{R})$ 的一个子域, 其中的 x, y, z, t 为任意实数. 并且这个子域同构于在前一个习题中定义的四元数域. 证明, 看作实向量空间, 这个域具有由满足下列条件的四个元素 e, i, j, k 组成的一个基:

$$\begin{aligned} &e^2 = e, \quad i^2 = j^2 = k^2 = -e,\\ &ei = ie = i, \quad ej = je = j, \quad ek = ke = k,\\ &ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j. \end{aligned}$$

如果允许变量 x, y, z, t 取任意复数值, 还得到一个域吗?

12. 基础环是 $\mathbf{C}$, 计算以下矩阵的逆矩阵:

$$\begin{pmatrix} 1 & 1+i & -i \\ 0 & i & 1-2i \\ 1 & 1 & i \end{pmatrix}.$$

13. 考虑 §§12, 13, 14 的习题 11 的矩阵 $\mathrm{U}(t)$. 进行尽可能少的计算求它的逆矩阵.

14. 设 K 是一个交换环. 证明形如

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

$(x, y, z \in \mathrm{K})$ 的矩阵组成 GL(3,K) 的一个子群. 确定这个子群的中心.

¶15. 设 K 是一个交换环, 而 n 是一个整数. 证明元素在 K 内的形如

$$\begin{pmatrix} 1 & * & * & \cdots & * \\ 0 & 1 & * & \cdots & * \\ 0 & 0 & 1 & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

的 n 阶方阵 (其中 $*$ 表示 K 的任意元素) 组成 $\mathrm{GL}(n, \mathrm{K})$ 的一个子群, 确定它的中心.

16. 设 K 是一个域. 用 H 表示由 $\mathrm{GL}(n, \mathrm{K})$ 的对角矩阵组成的 $\mathrm{GL}(n, \mathrm{K})$ 的子群. 求 H 在 $\mathrm{GL}(n, \mathrm{K})$ 内的正规化子 (§7, 习题 13).

¶17. $\mathbf{Z}^2$ 的元素 (a, b) 和 (c, d) 组成 $\mathbf{Z}^2$ 的一个基, 必须并且只需 $ad - bc = +1$ 或 -1.

18. 设 K 是一个交换环, 而 I 是 K 的一个理想. 设 H 是元素在 K 内的所有矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 的集合, 该矩阵满足条件

$$ad - bc = 1, \quad a \equiv d \equiv 1 (\mathrm{mod}\ \mathrm{I}), \quad b \equiv c \equiv 0 (\mathrm{mod}\ \mathrm{I}).$$

证明 H 是 $\mathrm{GL}(2, \mathrm{K})$ 的一个正规子群.

¶¶19. 设 K 是 q 个元素的有限域. 计算群 $\mathrm{GL}(n, \mathrm{K})$ 的元素个数.

¶20. 对于所有整数 $n \geqslant 1$, 用 G_n 表示矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 的集合, 其中

$$a, b, c, d \in \mathbf{Z}, \quad ad - bc = n,$$

并且令 $\mathrm{G}_1 = \mathrm{G}$.

a) 证明 G 是 $\mathrm{GL}(2, \mathbf{Z})$ 的一个子群. 对于 $n \geqslant 2$, G_n 也如此吗?

b) 证明如果 $\mathrm{X} \in \mathrm{G}_n$, 则对于任意 $\mathrm{U}, \mathrm{V} \in \mathrm{G}$ 有 $\mathrm{UXV} \in \mathrm{G}_n$.

c) 证明对于所有 $\mathrm{X} \in \mathrm{G}_n$, 存在 $\mathrm{U}, \mathrm{V} \in \mathrm{G}$, 使得 UXV 是对角的.

21. 设 K 是一个交换环. 在元素属于 K 并且满足条件 $ad - bc = 1$ 的矩阵的群 $\mathrm{SL}(2, \mathrm{K})$ 内, 考虑矩阵

$$x_+(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \quad x_-(t) = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix},$$

其中的 $t \in \mathrm{K}$. 用 α 表示两个符号 $+$ 或 $-$ 中的一个, 而用 $-\alpha$ 表示相反的符号, 对于 $t \neq 0$ 定义矩阵

$$w_\alpha(t) = x_\alpha(t) x_{-\alpha}(-1/t) x_\alpha(t), \quad h_\alpha(t) = w_\alpha(t) w_\alpha(1)^{-1}.$$

a) 计算这些矩阵, 并且证明矩阵 $x_+(t)$ 和 $x_-(t)$ 生成 $\mathrm{SL}(2, \mathrm{K})$.

b) 建立下列关系

(R1) $x_\alpha(t + u) = x_\alpha(t) x_\alpha(u)$, 对于 $t, u \in \mathrm{K}$;

(R2) $w_\alpha(t) x_\alpha(u) w_\alpha(t)^{-1} = x_{-\alpha}(-u/t^2)$, 对于 $t, u \in \mathrm{K}, t \neq 0$;

(R3) $h_\alpha(tu) = h_\alpha(t) h_\alpha(u)$, 对于 $t, u \in \mathrm{K}, t \neq 0, u \neq 0$.

c) 证明所有的矩阵 $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathrm{K})$ 可以用一种并且仅一种方式表示为形式 $g = h_+(t)x_+(u)$, 或者形式 $g = h_+(t)x_+(u)wx_+(v)$, 其中令 $w = w_+(1)$ (根据 $c = 0$ 或者 $c \neq 0$ 区分两种情形).

d) 称由其元素的 "交换子"$(x,y) = xyx^{-1}y^{-1}$ 生成的 G 的子群 G′ 为群 G 的导群. 证明 SL(2, K) 等于它的导群, 只要 K 至少含有四个元素.

¶¶ 22. 在 $\mathrm{SL}(2,\mathrm{K}) = \mathrm{G}$ 内考虑由 $x_+(t)$ 组成的子群 U 和由 $h_+(t)$ 组成的子群 H, 于是根据上题的问题 c), 有 $\mathrm{G} = \mathrm{HU} \cup \mathrm{HU}w\mathrm{U}$. 我们打算证明如果 K 至少含有四个元素, 那么群 G 仅包含 $\{e\}$ 和 G 以外的一个不变子群 M, 即由矩阵 1 和 -1 组成的子群 Z (它缩减为特征为 2 的中性元. 注意在所有情形, 这个子群是 G 的中心, 这一事实容易验证). 令 $\mathrm{B} = \mathrm{HU}$.

a) 证明 $\mathrm{B} \cap w\mathrm{B}w^{-1} = \mathrm{H}$, 并且由此推出如果 $\mathrm{M} \subset \mathrm{B}$, 则有 $\mathrm{M} \subset \mathrm{H}$.

b) 设 $\mathrm{M} \subset \mathrm{H}$. 假定 $h = h_+(t) \in \mathrm{M}$. 通过计算 h 和 $x_-(u)$ 的交换子证明 M 含有 $x_-(u - u/t^2)$, 由此推出 $\mathrm{M} \subset \mathrm{Z}$.

c) 假定 M 不包含于 B 内. 借助前一个习题的问题 c) 证明 $\mathrm{G} = \mathrm{MB}$, 并且由此推出 (§7, 习题 16) 商群 G/M 同构于 $\mathrm{B}/\mathrm{B} \cap \mathrm{M}$.

d) 假定 $\mathrm{Card(K)} \geqslant 4$, 根据前一个习题 $\mathrm{G} = \mathrm{G}'$. 证明群 G/M 等于它的导群. 由此和前一个问题推出: 如果 M 不包含于 B 内, 则 $\mathrm{M} = \mathrm{G}$ (注意到 $\mathrm{B}' \subset \mathrm{U}$ 和 $\mathrm{U}' = \{e\}$, 并且注意到 B 的仅有的使得 B/L 是自己的导群的不变子群 L 是 B 本身).

¶¶ 23. 在一个群 G 内, 考虑标记为 $\hat{x}_+(t)$ 和 $\hat{x}_-(t)$ 的两族元素, 它们依赖一个在一个给定的域 K 内变动的参数 t. 从 G 的元素 $\hat{x}_\alpha(t)$ 出发用习题 21 的公式定义 G 的元素 $\hat{w}_\alpha(t)$ 和 $\hat{h}_\alpha(t)$, 即令

$$\hat{w}_\alpha(t) = \hat{x}_\alpha(t)\hat{x}_{-\alpha}(-1/t)\hat{x}_\alpha(t), \quad \hat{h}_\alpha(t) = \hat{w}_\alpha(t)\hat{w}_\alpha(1)^{-1}.$$

假定习题 21 的关系满足

(R1) $\hat{x}_\alpha(t+u) = \hat{x}_\alpha(t)\hat{x}_\alpha(u)$,

(R2) $\hat{w}_\alpha(t)\hat{x}_\alpha(u)\hat{w}_\alpha(t)^{-1} = \hat{x}_{-\alpha}(-u/t^2)$.

a) 证明下列公式:

$$\hat{w}_\alpha(t)\hat{w}_\alpha(u)\hat{w}_\alpha(t)^{-1} = \hat{w}_{-\alpha}(-u/t^2), \quad \hat{w}_\alpha(t)\hat{h}_\alpha(u)\hat{w}_\alpha(t)^{-1} = \hat{h}_{-\alpha}(-u/t^2)\hat{h}_{-\alpha}(-1/t^2)^{-1},$$
$$\hat{w}_\alpha(t)\hat{x}_\alpha(u)\hat{w}_\alpha(t)^{-1} = \hat{x}_{-\alpha}(-u/t^2), \quad \hat{h}_\alpha(t)\hat{x}_\alpha(u)\hat{h}_\alpha(t)^{-1} = \hat{x}_\alpha(t^2u),$$
$$\hat{h}_\alpha(t)\hat{x}_\alpha(u)\hat{h}_\alpha(t)^{-1} = \hat{w}_\alpha(t^2u), \quad \hat{h}_\alpha(t)\hat{h}_\alpha(u)\hat{h}_\alpha(t)^{-1} = \hat{h}_\alpha(t^2u)\hat{h}_\alpha(t^2)^{-1},$$
$$\hat{w}_\alpha(-1/t) = \hat{w}_{-\alpha}(t), \quad \hat{w}_\alpha(t)^{-1} = \hat{w}_\alpha(-t),$$
$$\hat{h}_\alpha(-1/t) = \hat{h}_{-\alpha}(t), \quad \hat{h}_\alpha(t)\hat{h}_\alpha(-1/t) = \hat{h}_\alpha(-1),$$
$$\hat{w}_\alpha(1)\hat{h}_\alpha(t)\hat{w}_\alpha(1)^{-1} = \hat{h}_\alpha(1/t), \quad \hat{w}_\alpha(1)^{-2} = \hat{h}_\alpha(-1).$$

b) 设 U 是由 $\hat{x}_+(t)$ 生成的 G 的子群, 而 H 是由 $\hat{h}_+(t)$ 生成的 G 的子群. 令 $\hat{w} = \hat{w}_+(1)$ 和 $\mathrm{N} = \mathrm{H} \cup w\mathrm{H}$. 证明 N 是 G 的一个不变子群, 而 H 在 N 内是不变的. 证明我们有 $\hat{w}\mathrm{U}\hat{w} \subset \mathrm{UNU}$ (乘积 $u'nu''$ 的集合, 其中的 $u', u'' \in \mathrm{U}$, 而 $n \in \mathrm{N}$), 并且证明集合 $\mathrm{UH} \cup \mathrm{UH}\hat{w}\mathrm{U}$ 是由 $\hat{x}_\alpha(t)$ 生成的 G 的子群.

c) 重新取习题 21 的群 SL(2, K) 和这个群的元素 $x_\alpha(t), w_\alpha(t)$ 和 $h_\alpha(t)$, 利用习题 21 问题 c) 证明, 存在从 SL(2, K) 到 G 内的唯一的一个映射 π, 满足

$$\pi(h_+(t)x_+(u)) = \hat{h}_+(t)\hat{x}_+(u), \quad \pi(h_+(t)x_+(u)wx_+(v)) = \hat{h}_+(t)\hat{x}_+(u)\hat{w}\hat{x}_+(v).$$

证明: π 是一个同态, 必须并且只需关系

(R3) $\hat{h}_\alpha(tu) = \hat{h}_\alpha(t)\hat{h}_\alpha(u)$

满足. 换句话说, 为了构造从 SL(2, K) 到任何一个群 G 的一个同态, 必须并且只需给定 G 的满足关系 (R1), (R2) 和 (R3) 的元素 $x_+(t)$ 和 $x_-(t)$ (“用生成元和关系定义 SL(2, K)”).

¶¶24. 考虑一个交换域 K 上的群 $G = GL(n, K)$, 由对角矩阵组成的 G 的子群 T, 由其位于对角线下方的元素全是零的矩阵组成的子群 B, 以及其所有对角线元素等于 1 的 B 的矩阵组成的 B 的子群 U. 最后用 N 表示使得 $nTn^{-1} = T$ 的 $n \in G$ (T 在 G 内的正规化子) 的集合. 我们把 G 的元素等同于空间 K^n 的自同构, K^n 的典范基记作 $e_1, \cdots, e_n$.

a) 证明 $g \in N$ 当且仅当存在一个置换 $w \in \mathfrak{S}_n$ 和标量 $t_i \neq 0$, 使得对于 $1 \leqslant i \leqslant n$ 有 $g(e_i) = t_i e_{w(e_i)}$. 由此推出商群 $N/T = W$ 同构于对称群 $\mathfrak{S}_n$. 对于 $1 \leqslant i \leqslant n-1$, 设 $\omega_i \in G$ 是对换 e_i 为 e_{i+1}, 而对于 $j \neq i, i+1$ 令 e_j 不变的矩阵. 证明 **N** 是由 T 和诸 ω_i 生成的.

b) 对于 $i \neq j$ 和 $t \in K$, 用 $x_{ij}(t)$ 表示由下列公式定义的 G 的元素:

$$x_{ij}(t)e_j = e_j + te_i, \quad x_{ij}(t)e_k = e_k, \text{ 如果 } k \neq j.$$

$x_{ij}(t)$ 的矩阵是什么? 证明对于任意 $t, u \in K$ 有 $x_{ij}(t+u) = x_{ij}(t) + x_{ij}(u)$, 并且对于给定的 i 和 j 和变动的 t, $x_{ij}(t)$ 组成 G 的子群 U_{ij}. 按照一般方式, 令 $(a, b) = aba^{-1}b^{-1}$, 证明有

$$(x_{ij}(t), x_{jk}(u)) = x_{ik}(tu), \quad \text{如果 } i, j, k \text{ 两两相异},$$
$$(x_{ij}(t), x_{kl}(u)) = 1, \quad \text{如果 } j \neq k \text{ 并且 } i \neq l.$$

证明 U 是由 $x_{i,i+1}(t)(1 \leqslant i \leqslant n-1, t \in K)$ 生成的. 对于 $n \in N$ 计算 $nx_{ij}(t)n^{-1}$.

c) 设 B′ 是由其位于对角线上方的元素全是零的矩阵组成的 G 的子群. 证明存在 $n \in N$, 使得 $B' = nBn^{-1}$, 并且 B′ 是由 T 和 $x_{i+1,i}(t)$ 生成的. 设 g 是 G 的任何一个元素, 证明存在一个 $b \in B$ 和一个 $b' \in B'$, 使得 $g = b'bg_1$, 其中的 g_1 的形式是

$$g_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix}.$$

由此推出 G 是由 B 和 B′ 生成的, 或同样由 B 和 N 生成的 (关于 n 进行归纳推理).

利用习题 21 的公式证明 G 的由 $x_{i,i+1}(t)$ 和 $x_{j+1,j}(t)$ 生成的子群是 SL(n, K), 这是使得 $\det(g) = 1$ 的 $g \in GL(n, K)$ 的集合 (这个习题假定读者熟悉行列式理论, 后面用不到).

d) 考虑子群 $B_i = B \cap \omega_i^{-1} B \omega_i$. 证明 B_i 在 B 内是不变的, 并且所有 $b \in B$ 可以用唯一的方式写成 $U_{i,i+1}$ 和 B_i 的各一个元素的乘积. 由此推出双陪集 $B\omega_i B$, 即 $b'\omega_i b''(b', b'' \in B)$ 的集合, 是右陪集 $B\omega_i x_{i,i+1}(t)(t \in K)$ 的并集.

e) 令 $n^{-1}\mathrm{U}_{i,i+1}n = \mathrm{U}_{jk}$. 证明集合 $\mathrm{B}\omega_i\mathrm{B}n\mathrm{B}$ (对于 $n \in \mathrm{N}$) 是双陪集 $\mathrm{B}\omega_i n x_{jk}(t)\mathrm{B}$ 的并集, 其中 t 变动. 由此推出有

$$\mathrm{B}\omega_i\mathrm{B}n\mathrm{B} = \mathrm{B}\omega_i n\mathrm{B},\ \text{如果}\ j < k;\quad \mathrm{B}\omega_i\mathrm{B}n\mathrm{B} = \mathrm{B}n\mathrm{B} \cup \mathrm{B}\omega_i n\mathrm{B},\ \text{如果}\ j > k$$

(利用以下事实, 在 $\mathrm{GL}(2,\mathrm{K})$ 内有 $x_{i+1,i}(t) \in \mathrm{B}\omega_i\mathrm{B}$, 如果 $t \neq 0$).

f) 设 G_0 是双陪集 $\mathrm{B}n\mathrm{B}$ 的并集 (这些双陪集的数目是有限的, 因为 N/T 是有限的). 利用 N 是由 T 和各个 ω_i 生成这一事实, 并且借助前一个问题, 证明对于所有 $n \in \mathrm{N}$ 有 $n\mathrm{G}_0 \subset \mathrm{G}_0$. 证明 G_0 是 G 的一个子群, 并且由此推出

$$\mathrm{G} = \bigcup \mathrm{B}n\mathrm{B}$$

(对于线性群的 Bruhat 定理).

¶25. 设 K 是有 q 个元素的一个域, 而 V 是 K 上的有限 n 维向量空间. 设 m 是介于 0 和 n 之间的一个整数.

a) 设 X_m 是 V 的线性无关族 $(x_1, \cdots, x_m)$ 的集合, 证明有

$$\mathrm{Card}(\mathrm{X}_m) = (q^n - 1)(q^n - q)\cdots(q^n - q^{m-1}).$$

(关于 n 做归纳推理.)

b) 证明 V 的自同构群 GL(V) 的阶由以下公式给出:

$$\mathrm{Card}(\mathrm{GL(V)}) = (q^n - 1)(q^n - q)\cdots(q^n - q^{n-1}) = q^{n^2}\prod_{i=1}^{n}\left(1 - \frac{1}{q^i}\right).$$

c) 设 $\mathrm{G}_{n,m}$ 是 V 的 m 维向量子空间的 (Grassman) 集合. 证明有

$$\mathrm{Card}(\mathrm{G}_{n,m}) = q^{n^2-m^2}\prod_{i=m-1}^{n}\left(1 - \frac{1}{q^i}\right).$$

26. 设 V 是有两个元素的域上的一个二维向量空间; 设 x, y, z 是 V 的三个非零元素.

a) 证明有 $x + x = y + y = z + z = 0$ 和 $x + y = z, y + z = x, z + x = y$.

b) 由此推出 V 的自同构群 GL(V) 同构于 x, y, z 的置换群.

27. 设 V 是有三个元素的域上的一个二维向量空间.

a) 证明 V 包含四个一维子空间, 记为 $\mathrm{D}_1, \mathrm{D}_2, \mathrm{D}_3, \mathrm{D}_4$.

b) V 的自同构群 GL(V) 的所有元素在 D_i 之间置换它们. 由此推出有一个同态

$$\varepsilon : \mathrm{GL(V)} \to \mathrm{S}_4,$$

其中的 S_4 是 $\{1, 2, 3, 4\}$ 的置换群. 证明 ε 的核是 $\{\pm 1\}$; 由此推出 ε 是满射的 (比较两个群的阶).

c) 设 SL(V) 是由行列式为 1 的元素组成的 GL(V) 子群. 证明 ε 定义从 SL(V) 到由偶置换组成的交错群 A_4 的一个同构. [这个习题假定读者熟悉行列式理论.]

§16 线性映射的转置

1. 模的对偶

设 L 是任意环 K 上的一个右模. 我们在 §12 第 4 小节例 3 讲到, 称所有从 L 到右 K-模 K 内的同态为 L 上的线性型, 换句话说, L 上的线性型是映射

$$f : \mathrm{L} \to \mathrm{K},$$

它对于任意 $x, y \in \mathrm{L}$ 和 $\alpha, \beta \in \mathrm{K}$ 满足

$$f(x\alpha + y\beta) = f(x)\alpha + f(y)\beta. \tag{1}$$

根据 §13, 这些线性型是交换群 $\mathrm{Hom}(\mathrm{L}, \mathrm{K})$ 的元素, L 上的两个线性型的和 $f + g$ 由函数 $f(x) + g(x)$ 定义.

我们就要看到 (利用 K 不仅是右 K-模而且也是左 K-模这一事实), 事实上不仅可以把集合 $\mathrm{Hom}(\mathrm{L}, \mathrm{K})$ 看作一个交换群, 而且也可以看作一个*左* K-模. 设 f 是 L 上的线性型, $\lambda \in \mathrm{K}$ 是标量, 考虑 L 上的由

$$g(x) = \lambda \cdot f(x)$$

定义的函数. 关系 (1) 左乘以 λ, 考虑到在一个环内的计算规则 (结合性, 分配性), 则得

$$g(x\alpha + y\beta) = g(x)\alpha + g(y)\beta,$$

这就证明了 g 仍然是 L 上的一个线性型. 自然我们把它记作

$$g = \lambda f,$$

这样我们就对 $\mathrm{Hom}(\mathrm{L}, \mathrm{K})$ 定义了第二个运算, 它就是以一个标量 "乘" 这个集合的元素. 配备了 §13 定义的加法和刚才定义的第二个运算, $\mathrm{Hom}(\mathrm{L}, \mathrm{K})$ 实际上就是一个左 K-模. 让读者作为习题仔细验证 (还可以利用以下事实: 设 E 是所有从 L 到 K 内的线性的或不是线性的映射的集合, 把 L 看作一个任意的集合, 而把 K 看作一个左 K-模, §10 的例 4 允许把 E 看作一个左 K-模. 这一点交代清楚了, 那么为了证明 $\mathrm{Hom}(\mathrm{L}, \mathrm{K})$ 也是一个左 K-模, 只需证明 $\mathrm{Hom}(\mathrm{L}, \mathrm{K})$ 是 E 的一个子模, 而这正是前面考虑的目的).

配备了我们刚定义的左 K-模结构的集合 $\mathrm{Hom}(\mathrm{L}, \mathrm{K})$ 称为**右 K-模 L 的对偶**, 通常把它记作

$$\mathrm{L}^*,$$

这比记号 $\mathrm{Hom}(\mathrm{L}, \mathrm{K})$ 更简洁.

如果谈论左 K-模 L, 同样定义它的对偶, 这将得到一个右 K-模.

我们提醒, 关于这个主题, 如果环 K 是交换的, 则 "右" 和 "左" 之间的区别将毫无意义, 而在实际中, K 几乎都是交换的.



注 1 设 f 是右 K-模 L 上的一个线性型. 为了证明 λf 仍是线性的, 还可以把它看作 f 跟从 K 到 K 内的映射 $\xi \to \lambda\xi$ 的复合. 于是只需指出这后一个映射是右 K-模 K 的一个自同态, 即满足

$$\lambda(\xi+\eta)=\lambda\xi+\lambda\eta,\quad \lambda(\xi\mu)=(\lambda\xi)\mu,$$

而这是显而易见的.

设 L 是一个右 K-模. 因为在 L 上的线性型的集合 L^* 上已经定义了左 K-模结构, 就可以把模理论的定义和定理应用到 L^* 上. 这里是其中一个应用, 给定 L 上的线性型 $f_1,\cdots,f_p$, 我们说 L 上的线性型 f 是 $f_1,\cdots,f_p$ 的*线性组合*, 如果存在标量 $\lambda_1,\cdots,\lambda_p\in\mathrm{K}$, 使得

$$f=\lambda_1 f_1+\cdots+\lambda_p f_p.$$

考虑到线性型上运算的定义, 此式的意义是

$$f(x)=\lambda_1 f_1(x)+\cdots+\lambda_p f_p(x)\quad \text{对于所有 } x\in\mathrm{L}.$$

同样, 称所有的组

$$(\lambda_1,\cdots,\lambda_p)\in\mathrm{K}^p$$

为 $f_1,\cdots,f_p$ 之间的一个*线性关系*, 如果在 L^* 内有

$$\lambda_1 f_1+\cdots+\lambda_p f_p=0,$$

即

$$\lambda_1 f_1(x)+\cdots+\lambda_p f_p(x)=0\quad \text{对于所有 } x\in\mathrm{L}.$$

等等.

2. 有限生成自由模的对偶

在实践中, 初学者不需要比下述定理更 "深刻的" 结果:

定理 1 *设 L 是一个有限生成的自由的右 K-模, 而 $(a_1,\cdots,a_n)$ 是 L 的一个基. 考虑由*

$$\theta(f)=(f(a_1),\cdots,f(a_n))$$

定义的映射 $\theta:\mathrm{L}^*\to\mathrm{K}^n$, 那么 θ 是左 K-模之间的一个同构, 并且 L^* 具有一个基 $(u_1,\cdots,u_n)$, 使得

$$u_i(a_j)=\begin{cases}1, & \text{如果 } i=j,\\ 0, & \text{如果 } i\neq j.\end{cases}$$

设 $\alpha_1, \cdots, \alpha_n$ 是 K 的任意元素; §12 的定理 3 指出存在唯一的一个 $f \in \mathrm{L}^*$, 使得对于 $1 \leqslant i \leqslant n$ 有 $f(a_i) = \alpha_i$, 即

$$\theta(f) = (\alpha_1, \cdots, \alpha_n).$$

因而映射 θ 是双射的. 为了证明它是一个同构, 剩下的是证它是线性的. 对于 $f, g \in \mathrm{L}^*$, 令 $f + g = h$, 则有

$$\begin{aligned}\theta(f+g) &= (h(a_1), \cdots, h(a_n)) = (f(a_1) + g(a_1), \cdots, f(a_n) + g(a_n)) \\ &= (f(a_1), \cdots, f(a_n)) + (g(a_1), \cdots, g(a_n)) = \theta(f) + \theta(g);\end{aligned}$$

此外, f 换成 λf, 显然元素 $f(a_i)$ 左乘以 λ, 由此得到等式

$$\theta(\lambda f) = \lambda \cdot \theta(f),$$

从而 θ 是线性的.

为了结束定理的证明, 留下要证明具有所指出的性质的 L^* 的基 $(u_i)_{1 \leqslant i \leqslant n}$ 的存在性. 加在 u_i 上的条件意味着

$$\begin{gathered}\theta(u_1) = (1, 0, 0, \cdots, 0), \\ \theta(u_2) = (0, 1, 0, \cdots, 0), \\ \cdots\cdots\cdots\cdots \\ \theta(u_n) = (0, 0, 0, \cdots, 1),\end{gathered}$$

换句话说, θ 映射 L^* 的基 $(u_i)_{1 \leqslant i \leqslant n}$ 到 K^n 的典范基上. 由于 θ 是模的一个同构, u_i 的存在性是显然的: 只需取 K^n 的典范基在映射 θ^{-1} 下的像, 这就结束了证明.

我们注意到对于 L 的所有元素

$$x = a_1 \xi_1 + \cdots + a_n \xi_n,$$

有关系

$$u_i(x) = u_i(a_1)\xi_1 + \cdots + u_i(a_n)\xi_n = \xi_i.$$

也就是说, u_i 正是模关于 L 的基 $(a_i)_{1 \leqslant i \leqslant n}$ 的坐标函数 (§11, 第 4 小节). 对于所有 $f \in \mathrm{L}^*$, 设 $\alpha_1, \cdots, \alpha_n$ 是 f 关于基 $(u_i)_{1 \leqslant i \leqslant n}$ 的坐标, 那么关系

$$f = \alpha_1 u_1 + \cdots + \alpha_n u_n$$

写成

$$f(x) = \alpha_1 \xi_1 + \cdots + \alpha_n \xi_n \quad \text{对于所有 } x \in \mathrm{L},$$

因此得到

$$\alpha_i = f(a_i), \quad 1 \leqslant i \leqslant n.$$

换言之, f 关于 L^* 的基 (u_i) 的坐标正是在 §12 例 3 中定义的 f 关于 L 的基 (a_i) 的系数.

定理 1 表明, L 的所有基 (a_i) 对应 L^* 的一个基 (u_i), 我们说 (u_i) 是 L 的基(a_i) 的对偶基.

3. 模的二次对偶

设 L 是环 K 上的右模, 在它上面我们捆绑了一个左模 L^*, L^* 自己也具有一个对偶, 记作

$$L^{**} = (L^*)^*,$$

并且称为 L 的**二次对偶**. 像 L 一样, 这是一个右 K-模. 同样定义**三次对偶** $L^{***} = (L^{**})^*$, **四次对偶** $L^{****} = (L^{***})^*$, 如此无限进行下去.

在所有情形, 我们都可以用下列方式定义一个从 L 到 L^{**} 的**典范映射**: 设 x 是 L 的一个 "固定" 元素, 考虑由

$$u(f) = f(x) \quad \text{对于任意 } f \in L^*$$

定义的映射 $u: L^* \to K$, 它使得 L 上的每个线性型 f 对应它在 L 的点 x 的值. 映射 u 是线性的, 即对于任意 $f, g \in L^*$ 和数量 $\alpha, \beta \in K$, u 满足

$$u(\alpha f + \beta g) = \alpha \cdot u(f) + \beta \cdot u(g).$$

令

$$\alpha f + \beta g = h,$$

则上述关系可以写成

$$h(x) = \alpha f(x) + \beta g(x),$$

而在这种形式下, 该关系就纯粹地并且简单地归结为 L^* 的元素 $h = \alpha f + \beta g$ 的定义.

这样一来, u 是 L^* 上的一个线性型, 即 $u \in L^{**}$, 并且用这种方式, 我们使每个 $x \in L$, 对应了一个 $u \in L^{**}$. 由此得到从 L 到 L^{**} 内的一个映射, 定义这个映射是从 L 到它的二次对偶内的一个典范映射.

此外这个映射是线性的. 事实上, 设 $x, y \in L, \alpha, \beta \in K$, 并且令 $z = x\alpha + y\beta$; 还设 $u, v, w \in L^{**}$ 是 x, y, z 在典范映射下的像. 所有的事情都归结为证明关系

$$w = u\alpha + v\beta.$$

但由于 u, v, w 是 L^* 上的一个线性型, 上式意即

$$w(f) = u(f)\alpha + v(f)\beta \quad \text{对于任意 } f \in L^*;$$

而根据从 L 到 L^{**} 内的典范映射的定义, 我们有

$$u(f) = f(x), \quad v(f) = f(y), \quad w(f) = f(z),$$

因此一切都归结为证明

$$f(z) = f(x)\alpha + f(y)\beta,$$

用 z 的值代入, 此式变为

$$f(x\alpha + y\beta) = f(x)\alpha + f(y)\beta \quad \text{对于任意 } f \in L^*;$$

这个等式正是定义 L 上的线性型的等式.

定理 2 设 L 是有限生成的自由的模, 则典范映射

$$j : L \to L^{**}$$

是 L 到它的二次对偶的同构.

为了确立定理的正确性, 剩下要做的事情是证明: 如果 L 是有限生成的自由的, 则典范映射

$$j : L \to L^{**}$$

是双射. 设 (a_i) 是 L 的一个基, $(f_i)_{1\leqslant i\leqslant n}$ 是 L^* 的对偶基, 令

$$u_i = j(a_i) \in L^{**},$$

显然只需指出 u_i 组成 L^{**} 的一个基: 事实上, 我们要指出它们组成模 L^* 的基 (f_i) 在 L^{**} 内的对偶基, 即

$$u_i(f_j) = \begin{cases} 1, & \text{如果 } i = j, \\ 0, & \text{如果 } i \neq j. \end{cases}$$

按照从 L 到 L^{**} 的典范映射的定义, 我们有

$$u_i(f) = f(a_i) \quad \text{对于所有 } f \in L^*,$$

因此有 $u_i(f_j) = f_j(a_i)$. 而由于 f_j 组成 L 的基 (a_i) 的对偶基, 故有关系

$$f_j(a_i) = \begin{cases} 1, & \text{如果 } i = j, \\ 0, & \text{如果 } i \neq j. \end{cases}$$

这就提供了要找的关系, 并且完成了证明.

推论 1 设 L 是一个有限生成的自由的右 K-模, 而 u 是对偶模 L^* 上的一个线性型. 则存在唯一的一个 $x \in L$, **使得**

$$u(f) = f(x) \quad \text{对于所有 } f \in L^*.$$

事实上, 存在唯一的 $x \in \mathrm{L}$, 使得 $u = j(x)$, 这里 j 表示从 L 到它的二次对偶上的典范映射.

推论 2 L 是一个有限生成的自由的右 K-模, 而 $(f_1, \cdots, f_n)$ 是对偶模 L^* 的一个基. 那么对于任意 $\beta_1, \cdots, \beta_n \in \mathrm{K}$, 存在唯一的一个 $x \in \mathrm{L}$, 使得

$$f_i(x) = \beta_i \quad (1 \leqslant i \leqslant n).$$

事实上, 由于诸 f_i 组成 L^* 的一个基, 存在唯一的一个 L^* 上的线性型 u, 使得对于所有的 i 有 $u(f_i) = \beta_i$ (§12, 定理 3), 而 u 完全确定一个 $x \in \mathrm{L}$, 使得对于所有 $f \in \mathrm{L}^*$ 有 $u(f) = f(x)$, 推论证明完毕.

4. 同态的转置

设 L 和 M 是两个右 K-模, 而 $f: \mathrm{L} \to \mathrm{M}$ 是一个同态. 设 u 是 M 上的一个线性型, 那么复合映射 $u \circ f$ 显然是 (§12, 定理 1) L 上的一个线性型. 这样, 令

$${}^t f(u) = u \circ f \quad \text{对于所有 } u \in \mathrm{M}^*$$

就定义了一个映射

$${}^t f: \mathrm{M}^* \to \mathrm{L}^*,$$

映射 ${}^t f$ 称为**同态 f 的转置**.

这个映射跟 f 一样也是一个同态. 事实上, 设 $u, v \in \mathrm{M}^*$, 根据 §14 的定理 1 我们有

$${}^t f(u+v) = (u+v) \circ f = u \circ f + v \circ f,$$

因此有

$${}^t f(u+v) = (u+v) \circ f = {}^t f(u) + {}^t f(v).$$

同样, 对于 $\lambda \in \mathrm{K}$, 把 K 内的映射 $\xi \to \lambda\xi$ 记作 h_λ, 那么对于所有 $u \in \mathrm{M}^*$, 我们有

$${}^t f(\lambda u) = {}^t f(h_\lambda \circ u) = h_\lambda \circ u \circ f = h_\lambda \circ {}^t f(u) = \lambda \cdot {}^t f(u),$$

这就证明了转置 ${}^t f$ 是线性的.



注 2 刚给出的证明初学者看起来似乎是不可理解的, 原因是它的纯机械外表. 当然, 强烈建议读者重新翻译成清晰的语言 (一切都归结为一个模的对偶模的结构的定义). 然而, 也应当注意当陈述了一个定理 (例如 §14 的定理 1) 时, 我们总希望有机会利用它! 这个注同样适用于下一个结果的证明.

定理 3 从一个同态过渡到它的转置的运算具有下列性质:

a) 设 L, M 是两个右 K-模, 而 $f, g: \mathrm{L} \to \mathrm{M}$ 是两个同态, 则有

$${}^t(f+g) = {}^t f + {}^t g.$$

b) 设 L, M, N 是三个右 K-模, 而 $f: \mathrm{L} \to \mathrm{M}$ 和 $g: \mathrm{M} \to \mathrm{N}$ 是两个同态, 则有

$$^t(g \circ f) = {}^tf \circ {}^tg.$$

c) 设 L 是一个右 K-模, L 的自同构 (对应的, 恒等同构) 的转置是 L^* 的自同构 (对应的, 恒等自同构).

为了证明 a), 令 $h = f + g$. 对于 $u \in \mathrm{M}^*$ 我们有

$$^th(u) = u \circ h = u \circ (f+g) = u \circ f + u \circ g = {}^tf(u) + {}^tg(u),$$

这就证明了所宣布的等式 ${}^th = {}^tf + {}^tg$.

为了证明 b), 令 $h = g \circ f$. 对于 $u \in \mathrm{N}^*$ 有

$$^th(u) = u \circ h = u \circ (g \circ f) = (u \circ g) \circ f = {}^tg(u) \circ f = {}^tf[{}^tg(u)],$$

故得 ${}^th = {}^tf^t \circ {}^tg$, 这就证明了 b).

为了证明 c), 首先证明如果 $f: \mathrm{L} \to \mathrm{L}$ 是恒等映射, 那么它的转置也是恒等映射. 事实上, 对于 L 上的所有线性型 u, 由于 f 是恒等映射我们有

$$^tf(u) = u \circ f = u,$$

由此得到我们的结论. 再设 f 是 L 的一个自同构, 那么有一个 L 的同态 g, 使得

$$f \circ g = g \circ f = j_{\mathrm{L}},$$

因此有

$$^tg \circ {}^tf = {}^tf \circ {}^tg = {}^t(j_{\mathrm{L}}) = j_{\mathrm{L}^*},$$

从而 ${}^tf$ 是 L^* 的一个自同构. 此外像前面的计算所指出的, 对于 L 的所有自同构我们有

$$^t(f^{-1}) = ({}^tf)^{-1}.$$

注 3 留心不要把定理 3 的断言 b) 换成 "显然的" 却是错误的公式 (甚至是无意义的)



$$^t(f \circ g) = {}^tf \circ {}^tg.$$

5. 矩阵的转置

设 $f: \mathrm{L} \to \mathrm{M}$ 是有限生成自由右 K-模的一个同态. 选择 L 的一个基 $(a_1, \cdots, a_q)$ 和 M 的一个基 $(b_1, \cdots, b_p)$. 用 $(u_1, \cdots, u_q)$ 表示 L 的基 $(a_1, \cdots, a_q)$ 的对偶基, 用 $(v_1, \cdots, v_p)$ 表示 M 的基 $(b_1, \cdots, b_p)$ 的对偶基. 关于基 (a_j) 和 (b_i) 同态

$$f: \mathrm{L} \to \mathrm{M}$$

用形如

$$(\alpha_{ij})_{1\leqslant i\leqslant p,1\leqslant j\leqslant q}$$

的矩阵表示; 而关于基 (v_i) 和 (u_j), 转置同态

$${}^t f : \mathrm{M}^* \to \mathrm{L}^*$$

用形如

$$(\beta_{ji})_{1\leqslant i\leqslant p,1\leqslant j\leqslant q}$$

的矩阵表示. 我们打算用 f 的矩阵计算这个矩阵.

根据 ${}^t f$ 的矩阵定义, 我们有关系

$${}^t f(v_i) = \beta_{1i}u_1 + \cdots + \beta_{qi}u_q \quad (1 \leqslant i \leqslant p),$$

即

$$v_i \circ f = \beta_{1i}u_1 + \cdots + \beta_{qi}u_q,$$

这表示对于所有 $x \in \mathrm{L}$ 有

$$v_i[f(x)] = \beta_{1i}u_1(x) + \cdots + \beta_{qi}u_q(x). \tag{2}$$

令

$$x = a_1\xi_1 + \cdots + a_q\xi_q, \quad f(x) = b_1\eta_1 + \cdots + b_p\eta_p,$$

根据第 2 小节有关系

$$u_j(x) = \xi_j, \quad v_i(f(x)) = \eta_i.$$

因此关系 (2) 等价于

$$\eta_i = \beta_{1i}\xi_1 + \cdots + \beta_{qi}\xi_q,$$

而根据 f 的矩阵的定义还有关系

$$\eta_i = \alpha_{i1}\xi_1 + \cdots + \alpha_{iq}\xi_q,$$

比较两个关系的系数即得

$$\beta_{ji} = \alpha_{ij} \quad \text{对于 } 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant q.$$

这个结果引导我们引进以下定义. 给定一个元素在 K 内的矩阵

$$\mathrm{A} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1q} \\ \vdots & & \vdots \\ \alpha_{p1} & \cdots & \alpha_{pq} \end{pmatrix},$$

称置换矩阵 A 的行和列所得的矩阵

$$
{}^t\mathrm{A} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{p1} \\ \vdots & & \vdots \\ \beta_{1q} & \cdots & \alpha_{pq} \end{pmatrix}
$$

为 A **的转置**.

¶注 4 从 A 到 ^{t}A 的过渡显然对应着前面的从 f 到 ${}^tf$ 的过渡. 而 f 是右 K-模的一个同态, ${}^tf$ 则是左 K-模的一个同态, 这里左 K-模即是右 K°-模而 K° 是 K 的反环 (§10, 第 4 小节). 由于当 K 非交换时矩阵的计算适应于右模, 我们应当把元素在 K 内的一个矩阵 A 的转置 ^{t}A 看作是元素在反环 K° 内的一个矩阵. 我们将这样处理, 但在后面不再重新提及.



当然, 这些考虑当环 K 交换时是多余的.

定理 4 设 K 是一个环.

a) 给定两个元素在 K 内的矩阵 A 和 B, 只要和 A+B 有定义就有

$$
{}^t(\mathrm{A}+\mathrm{B}) = {}^t\mathrm{A} + {}^t\mathrm{B}.
$$

b) 给定元素在 K 内的矩阵 A 和 B, 只要乘积 AB 有定义就有关系(*)

$$
{}^t(\mathrm{AB}) = {}^t\mathrm{B}\cdot{}^t\mathrm{A}.
$$

c) 设 A 是一个元素在 K 内的方阵, A 是可逆的, 必须并且只需 ^{t}A 是可逆的.

d) 对于所有元素在 K 内的矩阵 A 有

$$
{}^t({}^t\mathrm{A}) = \mathrm{A}.
$$

为了证明断言 a), 令

$$
\mathrm{A} = (\alpha_{ij}), \quad \mathrm{B} = (\beta_{ij}),
$$

那么 $\mathrm{A}+\mathrm{B} = (\alpha_{ij}+\beta_{ij})$, 因此有

$$
{}^t(\mathrm{A}+\mathrm{B}) = (\alpha_{ji}+\beta_{ji}) = (\alpha_{ji}) + (\beta_{ji}) = {}^t\mathrm{A} + {}^t\mathrm{B}.
$$

为了证明断言 b), 令

$$
\mathrm{A} = (\alpha_{ij})_{1\leqslant i\leqslant p, 1\leqslant j\leqslant q}, \quad \mathrm{B} = (\beta_{jk})_{1\leqslant j\leqslant q, 1\leqslant k\leqslant r},
$$

矩阵

$$
\mathrm{AB} = (\gamma_{ik})_{1\leqslant i\leqslant p, 1\leqslant k\leqslant r}
$$

(*) 与注 4 相一致, 乘积 ${}^t\mathrm{B}\cdot{}^t\mathrm{A}$ 应当在 K 的反模 K° 中计算 (当 K 交换时就是在 K 内计算).

由关系

$$\gamma_{ik}=\sum_j \alpha_{ij}\beta_{jk} \tag{3}$$

给定 (§14, 第 2 小节), 乘积 $\alpha_{ij}\beta_{jk}$ 当然在给定的环 K 内计算. 此外我们有

$$^t\mathrm{A}=(\alpha'_{ji})_{1\leqslant j\leqslant q,1\leqslant i\leqslant p},\quad \text{其中 } \alpha'_{ji}=\alpha_{ij},$$
$$^t\mathrm{B}=(\beta'_{kj})_{1\leqslant k\leqslant r,1\leqslant j\leqslant q},\quad \text{其中 } \beta'_{kj}=\beta_{jk}.$$

令

$$^t\mathrm{B}\cdot{}^t\mathrm{A}=(\gamma'_{ki})_{1\leqslant k\leqslant r,1\leqslant i\leqslant p},$$

则有

$$\gamma'_{ki}=\sum_j \beta'_{kj}\alpha'_{ji},$$

其中乘积 $\beta'_{kj}\alpha'_{ji}$ 在 K 的反环 K° 内计算. 在 K 内计算此乘积, 并且考虑到关系 $\alpha'_{ji}=\alpha_{ij}$, $\beta'_{kj}=\beta_{jk}$, 则由 (3) 得到

$$\gamma'_{ki}=\sum_j \alpha'_{ji}\beta'_{kj}=\sum_j \alpha_{ij}\beta_{jk}=\gamma_{ik},$$

这就证明了出现在 b) 中的关系.

断言 d) 是平凡的.

剩下要证的是断言 c). 设 $\mathrm{A}\in\mathrm{M}_n(\mathrm{K})$, 假定 A 是可逆的, 并且设 B 是它的逆. 从关系 $\mathrm{AB}=\mathrm{BA}=1_n$ 得到 $^t\mathrm{B}\cdot{}^t\mathrm{A}={}^t\mathrm{A}\cdot{}^t\mathrm{B}={}^t(1_n)$. 显然有

$$^t(1_n)=1_n,$$

故 $^t\mathrm{A}$ 是可逆的, 并且所做的计算说明

$$(^t\mathrm{A})^{-1}={}^t(\mathrm{A}^{-1}).$$

反之, 设 $^t\mathrm{A}$ 是可逆的, 则刚证明的结果说明 $^t(^t\mathrm{A})$ 是可逆的, 即 A 是可逆的. 定理 4 至此完全证明.



注 5 显然可以从定理 3 的类似断言推导出定理 4 的断言 a) 和 b), 留给读者作为习题细心证明之.

§16 习题

1. 考虑 $\mathbf{R}^3$ 上的线性型

$$2x-y+3z,\quad 3x-5y+z,\quad 4x-7y+z,$$

它们组成 $\mathbf{R}^3$ 的对偶基吗?

2. 证明线性型

$$x+2y+z,\quad 2x+3y+3z,\quad 3x+7y+z$$

组成 $\mathbf{R}^3$ 的一个对偶基, 并且求与这个基对偶的 $\mathbf{R}^3$ 的基.

3. 设 K 是一个环, 而 $f_1,\cdots,f_n$ 是右 K-模 K^n 上的线性型. 这些线性型组成 K^n 的基的一个对偶基, 必须并只需存在向量 $x_1,\cdots,x_n\in\mathrm{K}^n$, 使得

$$f_i(x_j)=\begin{cases}1, & \text{如果 } i=j,\\ 0, & \text{如果 } i\neq j.\end{cases}$$

4. 设 K 是一个交换环, 而 U 是其元素在 K 内的一个 n 阶方阵.

a) 证明关系

$${}^t\mathrm{U}\cdot\mathrm{U}=1_n,\quad \mathrm{U}\cdot{}^t\mathrm{U}=1_n$$

是等价的.

b) 证明满足上述条件的矩阵 $\mathrm{U}\in\mathrm{M}_n(\mathrm{K})$ 组成 $\mathrm{GL}(n,\mathrm{K})$ 的一个子群 (环 K 上的 n **个变量的正交矩阵群**).

c) 称一个 $\mathrm{S}\in\mathrm{M}_n(\mathrm{K})$ 是对称的, 如果 ${}^t\mathrm{S}=\mathrm{S}$. 设 X 和 Y 是两个对称矩阵, XY 是对称的, 必须并且只需 $\mathrm{XY}=\mathrm{YX}$.

d) 证明 (取 $\mathrm{K}=\mathbf{Q}$ 或任何包含 $\mathbf{Q}$ 的域) 矩阵

$$\begin{pmatrix}\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2}\end{pmatrix}$$

同时是正交的和对称的.

e) 求所有元素为有理整数的三阶正交矩阵.

5. 设 $\mathrm{M}_n(\mathrm{K})$ 是任意环 K 上的 n 阶方阵的环, 把 $\mathrm{M}_n(\mathrm{K})$ 看作右 K-模. 证明对于 $\mathrm{M}_n(\mathrm{K})$ 上的所有的线性型 f, 存在唯一的一个矩阵 $\mathrm{A}\in\mathrm{M}_n(\mathrm{K})$, 使得

$$f(\mathrm{X})=\mathrm{Tr}(\mathrm{AX})\quad \text{对于所有 } \mathrm{X}\in\mathrm{M}_n(\mathrm{K})$$

(对于符号 Tr 的定义参见 §12 的习题 8). 对于任何 $\mathrm{X},\mathrm{Y}\in\mathrm{M}_n(\mathrm{K})$ 有

$$f(\mathrm{XY})=f(\mathrm{YX}),$$

当 K 可交换时, 必须并且只需矩阵 A 正比于 1_n.

§17 子模的和

1. 两个子模的和

设 K 是一个环, 而 L 是一个左 K-模, M 和 N 是 L 的两个子模. 称由 $\mathrm{M}\cup\mathrm{N}$ 生成的 L 的子模, 即同时包含 M 和 N 的 L 的最小子模, 为 **M 和 N 的和**.

容易给出这个子模的明晰结构: 它是这样的 $z \in \mathrm{L}$ 的集合, 对于它存在一个 $x \in \mathrm{M}$ 和一个 $y \in \mathrm{N}$, 使得

$$z = x + y.$$

事实上, 显然 L 的每一个包含 M 和 N 的子模含有具有上述性质的所有向量 z. 于是只需证明这些向量组成包含 M 和 N 的一个子模 P. 显然 P 包含 M 和 N (在关系 $z = x + y$ 中令 $x = 0$ 或 $y = 0$). 另一方面, 如果

$$z' = x' + y' \quad (x' \in \mathrm{M}, y' \in \mathrm{N}),$$
$$z'' = x'' + y'' \quad (x'' \in \mathrm{M}, y'' \in \mathrm{N})$$

是 P 的两个元素, 那么对于任何数量 λ 和 μ, 我们有

$$\lambda z' + \mu z'' = x + y,$$

其中

$$x = \lambda x' + \mu x'' \in \mathrm{M}, \quad y = \lambda y' + \mu y'' \in \mathrm{N},$$

这就证明了 P 是一个子模.

证明了上述结论, 我们用记号(*)

$$\mathrm{M} + \mathrm{N}$$

表示子模 M 和 N 的和

例 1 取 $\mathrm{K} = \mathbf{R}$, 取起点为给定的点 O 的通常空间的所有向量作为 L. 取过 O 的两条直线作为 M 和 N, 如果这两条直线是不同的, 那么 $\mathrm{M} + \mathrm{N}$ 是由这两条直线生成的平面, 如果它们是重合的, 那么 $\mathrm{M} + \mathrm{N} = \mathrm{M} = \mathrm{N}$.

前面的定义和构造直接推广到任意个数的子模. 设 $\mathrm{M}_1, \cdots, \mathrm{M}_p$ 是 L 的子模的任意一个有限族, 用 P 表示这样的 $z \in \mathrm{L}$ 的集合, 存在 $x_1 \in \mathrm{M}_1, \cdots, x_p \in \mathrm{M}_p$, 使得

$$z = x_1 + \cdots + x_p, \tag{1}$$

那么 P 是包含 $\mathrm{M}_1, \cdots, \mathrm{M}_p$ 的 L 的最小子模.

首先, P 包含 M_i (在前面的关系中对于 $j \neq i$ 取 $x_j = 0$), 包含 $\mathrm{M}_1, \cdots, \mathrm{M}_p$ 的子模的必然含有向量 (1), 这些都是显然的. 剩下要做的是证明 P 是 L 的一个子模. 如果

$$z' = x'_1 + \cdots + x'_p, \quad z'' = x''_1 + \cdots + x''_p$$

(*) 更一般的, 如果 G 是一个乘法 (对应的, 加法) 群, 而 A 和 B 是 G 的子集, 则用 AB (对应的, $\mathrm{A} + \mathrm{B}$) 表示这样的元素 $z \in \mathrm{G}$ 的集合, 对于 z 存在 $x \in \mathrm{A}$ 和 $y \in \mathrm{B}$, 使得 $z = xy$ (对应的, $z = x + y$).

是 P 的元素, 对于任何 $\lambda', \lambda'' \in \mathrm{K}$ 有关系

$$\lambda' z' + \lambda'' z'' = x_1 + \cdots + x_p,$$

其中

$$x_i = \lambda' x_i' + \lambda'' x_i'' \in \mathrm{M}_i \quad (1 \leqslant i \leqslant p),$$

这就证明了所要的结果.

这里跟前面一样, 我们说 P 是**子模** $\mathrm{M}_1, \cdots, \mathrm{M}_p$ **的和**, 并且记作

$$\mathrm{M}_1 + \cdots + \mathrm{M}_p.$$

2. 模的直积

在前面的内容中, 模 $\mathrm{M}_1, \cdots, \mathrm{M}_p$ 是一个模 L 的子模. 现在要从左 K-模 $\mathrm{M}_1, \cdots, \mathrm{M}_p$ 出发, 不假定它们包含于同一个模内, 而要构造一个包含所有 M_i 的模, 当然这里对于同构的模不加区别.

为此, 考虑由族

$$x = (x_1, \cdots, x_p), \quad x_1 \in \mathrm{M}_1, \cdots, x_p \in \mathrm{M}_p$$

组成的乘积

$$\mathrm{L} = \mathrm{M}_1 \times \cdots \times \mathrm{M}_p.$$

我们要在 L 上定义一个左 K-模结构, 为此令

$$(x_1, \cdots, x_p) + (y_1, \cdots, y_p) = (x_1 + y_1, \cdots, x_p + y_p),$$
$$\lambda(x_1, \cdots, x_p) = (\lambda x_1, \cdots, \lambda x_p).$$

用这种方式我们得到 L 上的模结构这个事实可以直接验证, 验证的任务留给读者细心去做.

当 $\mathrm{M}_1 = \cdots = \mathrm{M}_p = \mathrm{K}$ 时显然重新得到 K-模 K^p.

当 $\mathrm{K} = \mathbf{Z}$ 时, 重新回到在 §7 第 2 小节定义的交换群直积的概念.

在一般情形, 我们说配备了刚才定义的模结构的 $\mathrm{M}_1 \times \cdots \times \mathrm{M}_p$ 为 **模** $\mathrm{M}_1, \cdots, \mathrm{M}_p$ **的直积**.

显然对于 $1 \leqslant i \leqslant p$, 由 $\mathrm{pr}_i(x_1, \cdots, x_p) = x_i$ 定义的映射

$$\mathrm{pr}_i : \mathrm{M}_1 \times \cdots \times \mathrm{M}_p \to \mathrm{M}_i$$

是模的同态. 同样由

$$u_i(x) = (0, \cdots 0, x, 0, \cdots, 0)$$

定义的映射 $u_i : \mathrm{M}_i \to \mathrm{M}_1 \times \cdots \times \mathrm{M}_p$ 也是同态, 其中 $x \in \mathrm{M}_i$, 前面有 $i-1$ 个零.

事实上, 同态 u_i 是一个单射, 因此这是 M_i 到 $\mathrm{M}_1\times\cdots\times\mathrm{M}_p$ 的一个子模上的同构. 在实际中, 经常把 M_i 跟它在 u_i 下的像等同. 公式

$$(x_1,\cdots,x_p)=(x_1,0,\cdots,0)+(0,x_2,0,\cdots,0)+\cdots+(0,\cdots,0,x_n)$$

表明 $\mathrm{M}_1\times\cdots\times\mathrm{M}_p$ 的所有元素是子模 M_1 的一个元素, M_2 的一个元素, $\cdots$, M_p 的一个元素之和. 换句话说 $\mathrm{M}_1\times\cdots\times\mathrm{M}_p$ 是子模 $\mathrm{M}_1,\cdots,\mathrm{M}_p$ 的和.

3. 子模的直和

像在第 1 小节那样再取左 K-模 L 的子模 $\mathrm{M}_1,\cdots,\mathrm{M}_p$. 考虑由

$$f(x_1,\cdots,x_p)=x_1+\cdots+x_p$$

定义的映射

$$f:\mathrm{M}_1\times\cdots\times\mathrm{M}_p\to\mathrm{L}.$$

这显然是模的同态. 根据子模的一个族的和的定义本身得到

$$\mathrm{Im}(f)=\mathrm{M}_1+\cdots+\mathrm{M}_p.$$

当同态 f 是单射时, 就说子模 $\mathrm{M}_1,\cdots,\mathrm{M}_p$ 是线性无关的, 即所有 $x\in\mathrm{M}_1+\cdots+\mathrm{M}_p$ 以唯一方式写成形式 $x=x_1+\cdots+x_p$, 其中 $x_1\in\mathrm{M}_1,\cdots,x_p\in\mathrm{M}_p$. 同样归结为说 (§7, 定理 8) $\mathrm{Ker}(f)$ 缩减为 0, 就是说关系

$$x_1+\cdots+x_p=0,\quad x_1\in\mathrm{M}_1,\cdots,x_p\in\mathrm{M}_p$$

蕴含

$$x_1=0,\cdots,x_p=0.$$

定理 1 一个模 L 的两个子模 M 和 N 是线性无关的, 必须且只需 $\mathrm{M}\cap\mathrm{N}=\{0\}$.

设 $x\in\mathrm{M}\cap\mathrm{N}$, 那么 $x+(-x)=0$, 其中 $x\in\mathrm{M},-x\in\mathrm{N}$, 如果 M 和 N 是线性无关的, 故 $x=0$. 反之, 假定 $\mathrm{M}\cap\mathrm{N}=\{0\}$; 如果 $x\in\mathrm{M}$ 和 $y\in\mathrm{N}$ 满足 $x+y=0$, 则有 $x=-y\in\mathrm{M}\cap\mathrm{N}$, 故 $x=y=0$, 这就完成了证明.



注 1 考虑由

$$f(x,y)=x+y$$

给定的映射 $f:\mathrm{M}\times\mathrm{M}\to\mathrm{L}$. 前面的推理说明 f 的核由 $(x,-x)$ 组成, 其中 $x\in\mathrm{M}\cap\mathrm{N}$. 显然映射 $x\to(x,-x)$ 是从 $\mathrm{M}\cap\mathrm{N}$ 到 $\mathrm{Ker}(f)$ 的一个同构. 这个注对于计算子向量空间的和的维数十分重要 (§19, 第 7 小节).

当模 L 的子模 $\mathrm{M}_1,\cdots,\mathrm{M}_p$ 线性无关时, 称 $\mathrm{M}_1+\cdots+\mathrm{M}_p$ 是给定子模的**直和**, 用记号

$$\mathrm{M}_1\oplus\cdots\oplus\mathrm{M}_p$$

表示这个直和. 用符号 $\oplus$ 代替通常符号 $+$ 指明谈论的是线性无关的子模的和. 显然这时上面定义的映射 f 是从模 $\mathrm{M}_1 \times \cdots \times \mathrm{M}_p$ 到 $\mathrm{M}_1 \oplus \cdots \oplus \mathrm{M}_p$ 上的一个同构. 称这个同构为"典范的".

注 2 直积 $\mathrm{M}_1 \times \cdots \times \mathrm{M}_p$ 显然是在上一小节说的与 M_i $(1 \leqslant i \leqslant p)$ 等同的子模 $u_i(\mathrm{M}_i)$ 的直和.



注 3 设 $a_1, \cdots, a_p$ 是左 K-模 L 的元素, 而 $\mathrm{M}_1, \cdots, \mathrm{M}_p$ 是分别由 $a_1, \cdots, a_p$ 生成的子模. 那么 $\mathrm{M}_1 + \cdots + \mathrm{M}_p$ 就是由 $a_1, \cdots, a_p$ 生成的 L 的子模. 子模 $\mathrm{M}_1, \cdots, \mathrm{M}_p$ 是线性无关的, 当且仅当 $a_1, \cdots, a_p$ 是线性无关的; 而关系

$$\mathrm{L} = \mathrm{M}_1 \oplus \cdots \oplus \mathrm{M}_p$$

表明 $a_1, \cdots, a_p$ 组成 L 的一个基. 留给读者仔细验证这个断言.

设 L 是一个左 K-模, 而 M 是 L 的一个子模. 说 M 是 L 的一个**直和项**, 如果存在 L 的一个子模 N, 使得 L 是 M 和 N 的直和; 这时说 N 是 M 在 L 内的一个**补子模**. 在 §19 将看到, 如果 L 是一个域上的有限维向量空间, L 的所有子空间具有一个补子空间. 但是这个性质推广不到任意环.

例 2 取 $\mathrm{K} = \mathrm{L} = \mathbf{Z}$, $\mathrm{M} = p\mathbf{Z}$, $p \neq 0$. 设 $\mathrm{N} = q\mathbf{Z}$ 是 L 的非 0 子模, 那么有 $\mathrm{M} \cap \mathrm{N} \neq \{0\}$, 比如 $pq \in \mathrm{M} \cap \mathrm{N}$, 因此 (定理 1) M 在 $\mathbf{Z}$ 内不具有补 (当然除去情形 $\mathrm{M} = \mathrm{L}$ 和 $\mathrm{M} = \{0\}$).

4. 直和与投影

设 L 是一个左 K-模, 考虑 L 的一个**直和分解**, 即形如

$$\mathrm{L} = \mathrm{M}_1 \oplus \cdots \oplus \mathrm{M}_p$$

的一个关系, 其中诸 M_i 是 L 的线性无关的子模. 所有的 $x \in \mathrm{L}$ 以唯一的一种方式写成

$$x = x_1 + \cdots + x_p, \quad \text{其中对于 } 1 \leqslant i \leqslant p \text{ 有 } x_i \in \mathrm{M}_i,$$

由此可以令

$$x_i = v_i(x),$$

这里 v_i 是从 L 到 L 内的映射. 另外, 如果引进由

$$f(x_1, \cdots, x_p) = x_1 + \cdots + x_p$$

定义的同构和典范单射

$$j_i : \mathrm{M}_i \to \mathrm{L},$$

那么显然有

$$v_i = j_i \circ \mathrm{pr}_i \circ f^{-1}.$$

这就表明 v_i 是线性的 (读者也可以通过直接的计算证明这个结果).

自同态 v_i 的像显然是子模 M_i. 对于所有 $x \in \mathrm{L}$, $v_i(x)$ 是 M_i 的唯一向量, 使得 $x - v_i(x)$ 属于由 $\mathrm{M}_1, \cdots, \mathrm{M}_{i-1}, \mathrm{M}_{i+1}, \cdots, \mathrm{M}_p$ 生成的空间. 基于这个理由, 我们称 $v_i(x)$ 是 x 在 M_i 上的**投影**.

显然, 当且仅当 $x \in \mathrm{M}_i$, $v_i(x) = x$. 由于对于所有 $x \in \mathrm{L}$ 有 $v_i(x) \in \mathrm{M}_i$, 因此对于所有 $x \in \mathrm{L}$ 有 $v_i(v_i(x)) = v_i(x)$, 这就是说

$$v_i \circ v_i = v_i.$$

说模 L 的一个自同态 v 是一个**投影**, 如果它满足前面的条件, 即

$$v \circ v = v.$$

另外显然对于 $x \in \mathrm{M}_i, i \neq j$ 有 $v_j(x) = 0$, 于是对于所有 $x \in \mathrm{L}$ 和 $i \neq j$ 有 $v_j(v_i(x)) = 0$, 即

$$\text{如果 } i \neq j, \text{ 则 } v_j \circ v_i = 0.$$

最后, 对于所有 $x \in \mathrm{L}$ 有 $x = x_1 + \cdots + x_p = v_1(x) + \cdots + v_p(x)$, 这证明了

$$v_1 + \cdots + v_p = j_{\mathrm{L}},$$

这里 j_{L} 是从 L 到自身内的恒等映射.

这些性质具有一个逆命题:

定理 2　设 $v_1, \cdots, v_p$ 是模 L 的自同态, 满足下列关系:

$$v_j \circ v_i = \begin{cases} v_i, & \text{如果 } i = j, \\ 0, & \text{如果 } i \neq j, \end{cases}$$
$$v_1 + \cdots + v_p = j_{\mathrm{L}},$$

则 L 是子模 $\mathrm{M}_i = v_i(\mathrm{L})$ 的直和.

关系

$$x = v_1(x) + \cdots + v_p(x)$$

已经说明 L 是子模 M_i 的和. 现在考虑 $x_i \in \mathrm{M}_i$ $(1 \leqslant i \leqslant p)$, 它们满足

$$x_1 + \cdots + x_p = 0. \tag{2}$$

存在 $y_j \in \mathrm{L}$, 使得 $x_i = v_i(y_i)$, 于是有

$$v_i(x_j) = v_i[v_j(y_j)] = \begin{cases} 0, & \text{如果 } i \neq j, \\ v_i(y_i) = x_i, & \text{如果 } i = j. \end{cases}$$

把 v_i 作用在关系 (2) 上, 留下关系 $x_i = 0$, 这就表明子模 M_i 是线性无关的, 证明因此完成.

推论 设 M 是模 L 的一个子模. 以下性质是等价的:

(FD1) M 是 L 的直和项;

(FD2) 存在 L 的一个自同态, 使得

$$v \circ v = v, \quad v(\mathrm{L}) = \mathrm{M};$$

(FD3) 存在从 L 到 M 内的一个同态 q, 使得对于所有 $x \in \mathrm{M}$ 有 $q(x) = x$.

显然 (FD1) 蕴含 (FD2): 写出 $\mathrm{L} = \mathrm{M} \oplus \mathrm{N}$, 取 $v(x)$ 为 x 在 M 上的 "平行于 N" 的投影.

(FD2) 蕴含 (FD3): 由于 v 映射 L 到 M 上, 可以对于所有 $x \in \mathrm{L}$ 令 $q(x) = v(x)$, 从而定义一个从模 L 到模 M 内的同态 q (v 和 q 的唯一区别是, v 是从 L 到 L 内的一个映射, 而 q 是从 L 到 M 内的一个映射). 对于 $x \in \mathrm{M}$, 存在 $y \in \mathrm{L}$, 使得 $v(y) = x$, 于是有 $q(x) = v(x) = v(v(y)) = v(y) = x$.

最后证明 (FD3) 蕴含 (FD1). 设 N 是 q 的核. 我们有

$$\mathrm{M} \cap \mathrm{N} = \{0\},$$

这是因为, 如果 $x \in \mathrm{M} \cap \mathrm{N}$, 那么一方面有 $q(x) = x$, 另一方面有 $q(x) = 0$. 另外, 对于所有 $x \in \mathrm{L}$, 我们有 $q(q(x)) = q(x)$, 从而 $q(x - q(x)) = 0$, 故 $x - q(x) \in \mathrm{N}$. 写出 $x = q(x) + (x - q(x)) \in \mathrm{M} + \mathrm{N}$, 就发现 $\mathrm{L} = \mathrm{M} + \mathrm{N}$. 由于 $\mathrm{M} \cap \mathrm{N} = \{0\}$, 这就证明了 (定理 1) M 是在 L 内的一个直和项, 并且完成了证明.

§17 习题

1. 设 K 是一个环, L 是一个 K-模, 并且

$$\mathrm{L} = \mathrm{M}_1 \oplus \cdots \oplus \mathrm{M}_n$$

是 L 的子模直和分解.

a) 证明: 对于 L 的所有自同态 u, 存在一个并且仅一组同态

$$u_{ji} : \mathrm{M}_i \to \mathrm{M}_j \quad (1 \leqslant i, j \leqslant n),$$

使得有

$$u(x) = u_{1i} + \cdots + u_{ni} \quad \text{对于所有 } x \in \mathrm{M}_i \text{ 和所有 } i : 1 \leqslant i \leqslant n.$$

反之, 证明: 如果给定了这样的同态 u_{ji}, 则存在 L 的唯一的一个满足上述条件的自同态 u.

b) 令 L 的所有自同态 u 对应由问题 a) 定义的同态组成的 "矩阵"

$$\begin{pmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & & \vdots \\ u_{n1} & \cdots & u_{nn} \end{pmatrix}.$$

给定 L 的两个自同态 u 和 v, 怎样用 u 和 v 的矩阵计算 $v \circ u$ 的矩阵?

2. 设 r 是给定的 $\geqslant 1$ 的整数, 而整数 $r_1, \cdots, r_n$ 是 $\geqslant 1$ 的整数, 满足条件

$$r = r_1 + \cdots + r_n.$$

给定一个系数在一个任意的环 K 内的方阵

$$\mathrm{U} = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix},$$

用 $\mathrm{U}_{ij}(1 \leqslant i, j \leqslant n)$ 表示由矩阵 U 的元素 a_{pq} 组成的 (r_i 列 r_j 行的) 矩阵, 其中

$$r_1 + \cdots + r_{i-1} < p \leqslant r_1 + \cdots + r_i,$$

$$r_1 + \cdots + r_{j-1} < q \leqslant r_1 + \cdots + r_j,$$

这样按照显然的约定, 就可以把 U 写成

$$\mathrm{U} = \begin{pmatrix} \mathrm{U}_{11} & \cdots & \mathrm{U}_{1n} \\ \vdots & & \vdots \\ \mathrm{U}_{n1} & \cdots & \mathrm{U}_{nn} \end{pmatrix}.$$

设 V 是元素在 K 内的另一个 r 阶方阵, 且如同 U 那样分成块 V_{ij}, 而 $\mathrm{W} = \mathrm{VU}$ 是乘积矩阵. 证明组成 W 的 "块" W_{ij} 由类似于通常矩阵计算规则的公式

$$\mathrm{W}_{ij} = \mathrm{V}_{i1}\mathrm{U}_{1j} + \cdots + \mathrm{V}_{in}\mathrm{U}_{nj}$$

给出 (矩阵的**分块乘积公式**). 这个结果可以推广到一般矩阵吗?

3. 沿用前一个习题的记号, 证明矩阵 U 的转置由下式给定

$$^t\mathrm{U} = \begin{pmatrix} ^t\mathrm{U}_{11} & \cdots & ^t\mathrm{U}_{n1} \\ \vdots & & \vdots \\ ^t\mathrm{U}_{1n} & \cdots & ^t\mathrm{U}_{nn} \end{pmatrix}.$$

4. 设 K 是一个交换域. 考虑方阵

$$\mathrm{J} = \begin{pmatrix} 0 & -1_n \\ 1_n & 0 \end{pmatrix}$$

(其中的 0 是 n 行 n 列的矩阵), 求其元素在 K 内的 $2n$ 阶方阵

$$\mathrm{U} = \begin{pmatrix} \mathrm{A} & \mathrm{B} \\ \mathrm{C} & \mathrm{D} \end{pmatrix},$$

使得

$$^t\mathrm{U} \cdot \mathrm{J} \cdot \mathrm{U} = \mathrm{J}$$

(其中的 $\mathrm{A, B, C, D}$ 是其元素在 K 内的 n 阶方阵). 写出 $\mathrm{A, B, C, D}$ 应当满足的关系. 求 $\mathrm{C} = 0$ 时的矩阵 U.

5. 设 $\mathrm{S} = {}^t\mathrm{S}$ 是其元素在一个交换域 K 内的 n 阶方阵. 令

$$\mathrm{J} = \begin{pmatrix} 0 & 0 & 1_p \\ 0 & \mathrm{S} & 0 \\ 1_p & 0 & 0 \end{pmatrix}$$

(使得 J 是一个 $n + 2p$ 阶方阵). 在什么条件下矩阵

$$\mathrm{M} = \begin{pmatrix} \mathrm{U} & \mathrm{X} & \mathrm{Z} \\ 0 & \mathrm{V} & \mathrm{Y} \\ 0 & 0 & \mathrm{W} \end{pmatrix}$$

(其中 M 的块分解跟 J 有同样的形式) 满足关系

$${}^t\mathrm{M} \cdot \mathrm{J} \cdot \mathrm{M} = \mathrm{J}?$$

6. 设 K 是一个环, 设 $r \geqslant 1$ 是给定的整数, 而 $r_1 \geqslant 1, \cdots, r_n \geqslant 1$ 是整数, 令 $r = r_1 + \cdots + r_n$, 考虑矩阵

$$\mathrm{U} = \begin{pmatrix} \mathrm{U}_{11} & \cdots & \mathrm{U}_{1n} \\ \vdots & & \vdots \\ \mathrm{U}_{n1} & \cdots & \mathrm{U}_{nn} \end{pmatrix}$$

(其中的 U_{ij} 是 r_i 行 r_j 列的矩阵), 它们满足下列条件:

$$\mathrm{U}_{ij} = 0, \quad 如果\ i > j,$$
$$\mathrm{U}_{ii}\ 是可逆的 \quad (1 \leqslant i \leqslant n).$$

证明矩阵 U 组成一般线性群 $\mathrm{GL}(r, \mathrm{K})$ 的一个子群. 对于满足条件

$$\mathrm{U}_{ij} = 0, \quad 如果\ i < j,$$
$$\mathrm{U}_{ii} = 1 \qquad (1 \leqslant i \leqslant n)$$

的矩阵 U 解答同样的问题. 证明第二个子群在第一个子群内是不变的, 并且由幂幺矩阵组成.

7. 设 $\mathrm{L} = \mathrm{M}_1 \oplus \cdots \oplus \mathrm{M}_p$ 是环 K 上的一个左 K-模 L 的直和分解. 在 L 的对偶右 K-模 L^* 内, 对于每一个 $i (1 \leqslant i \leqslant p)$ 考虑 L 上的满足下列条件的线性型 f 的集合 M'_i:

$$f(\mathrm{M}_i) = 0 \quad 对于所有\ j \neq i.$$

证明 M'_i 是 L^* 的子模, 并且 $\mathrm{L}^* = \mathrm{M}'_1 \oplus \cdots \oplus \mathrm{M}'_p$.

8. 设 M 是环 K 上的一个左模, 而 M' 是 M 的一个子模. 假定商模 M/M' (§§10, 11, 习题 10) 是有限生成自由的. 证明这时 M' 是 M 内的一个直和项 (考虑由这样的元素生成的 M 的子模, 这些元素在典范映射下在 M/M' 内的像组成 M/M' 的一个基). 对于这个结果的一个重要应用, 见 §29, 习题 11, g).

¶9. 设 M 是环 K 上的一个左模, 而 a 是 M 的这样的一个元素, 它满足条件: $\lambda \in \mathrm{K}$, 且 $\lambda a = 0$ 蕴含 $\lambda = 0$. 由 a 生成的子模 Ka 是 M 内的一个直和项, 必须并且只需存在 M 上的一个这样的线性型 f, 它使得 $f(a) = 1$, 并且有

$$\mathrm{M} = \mathrm{K}a \oplus \mathrm{Ker}(f).$$

第四章 有限维向量空间

§18 至 §20 的目的是研究在任意域上的有限维向量空间, 特别是要引进维数这一基本概念, 并且推导在一个域上的线性方程组的最重要的性质.

虽然在初等实践中, 在分析里, 人们仅对于基础环是实数域和复数域的情况感兴趣, 但是在本章假定 $\mathrm{K} = \mathbf{R}$ 或者 $\mathrm{K} = \mathbf{C}$ 不会带来任何简化.

线性代数的经典陈述非常多地利用行列式理论, 但是人们能够跨越这一理论已经有 50 多年了, 这样得到的纯粹 "几何的" 陈述比建立在行列式理论上的陈述不仅更简单, 而且更一般, 因为后者假定基础的环或域是交换的. 事实上, 行列式理论 (将在更后面叙述) 的重要性在于提供线性无关的明晰的判别法和线性方程组的明晰的求解公式, 而对于建立 §19 和 §20 的存在性理论却没有任何作用.

线性代数的最重要的结果之一是 §19 的定理 13, 它在给定一个有限维向量空间 L 到另一个向量空间的同态 f 之后, 在 L 的维数与 f 的核和 f 的像的维数之间建立了一个简单的关系. 当分析这个定理的古典证明时, 我们坚信可以这样表述它, 使得它能够推广到一个任意环上的模. 利用这一般的结果可以十分简单地证明一些陈述, 这些陈述是所谓 Noether 环和主理想整环的 "重大" 理论的出发点. 这些结果是 §18 讨论的对象, 这节和对应习题的研究对于理解后续内容虽然原则上用处不大, 却能够使读者对于当前代数的最重要理论之一有一个概观的了解.

§18 有限性定理

1. 其核与像均为有限生成的同态(*)

本节的结果建立在以下定理的基础上:

定理 1 设 K 是一个环, L 和 M 是左 K-模, 而 f 是从 L 到 M 内的一个同态. 如果 $\mathrm{Ker}(f)$ 和 $\mathrm{Im}(f)$ 是有限生成的 K-模, 则 L 也是.
如果 $\mathrm{Ker}(f)$ 同构于 K^p, 而 $\mathrm{Im}(f)$ 同构于 K^q, 则 L 同构于 K^{p+q}.

假定 $\mathrm{Ker}(f)$ 由向量 $a_1, \cdots, a_p$ 生成, 而 $\mathrm{Im}(f)$ 由向量 $b_1, \cdots, b_q$ 生成. 在 L 中选取向量 $a_{p+1}, \cdots, a_{p+q}$, 使得

$$b_1 = f(a_{p+1}), \cdots, b_q = f(a_{p+q}).$$

为了证明定理 1, 只需证明一方面在所有情形下 $a_1, \cdots, a_{p+q}$ 生成 L, 而另一方面如果 $a_1, \cdots, a_p$ 组成 $\mathrm{Ker}(f)$ 的一个基, 而 $b_1, \cdots, b_q$ 组成 $\mathrm{Im}(f)$ 的一个基, 则 $p+q$ 个向量 $a_1, \cdots, a_{p+q}$ 组成 L 的一个基.

为了证明第一个断言, 考虑一个向量 $x \in \mathrm{L}$. 由于 $f(x)$ 属于由 $b_1, \cdots, b_q$ 生成的 M 的子模, 存在标量 $\eta_j (1 \leqslant j \leqslant q)$, 使得

$$f(x) = \eta_1 b_1 + \cdots + \eta_q b_q,$$

此式还可以写成

$$f(x) = \eta_1 f(a_{p+1}) + \cdots + \eta_q f(a_{p+q}) = f(\eta_1 a_{p+1} + \cdots + \eta_q a_{p+q}),$$

故有

$$x = \eta_1 a_{p+1} + \cdots + \eta_q a_{p+q} + y,$$

其中 $y \in \mathrm{Ker}(f)$. 但这时存在标量 $\xi_i (1 \leqslant i \leqslant p)$, 使得

$$y = \xi_1 a_1 + \cdots + \xi_p a_p,$$

把这个结果代入到前面的关系中, 即可发现 x 必然是向量 $a_1, \cdots, a_{p+q}$ 的线性组合.

为了证明第二个断言, 只需指出如果向量 $a_i (1 \leqslant i \leqslant p)$ 是线性无关的, 向量 $b_j (1 \leqslant j \leqslant q)$ 是线性无关的, 那么 $a_1, \cdots, a_{p+q}$ 也是线性无关的. 考虑形如

$$\lambda_1 a_1 + \cdots + \lambda_{p+q} a_{p+q} = 0$$

的关系, 把 f 作用在左端. 由于 $a_1, \cdots, a_p \in \mathrm{Ker}(f)$, 我们得到

$$\lambda_{p+1} f(a_{p+1}) + \cdots + \lambda_{p+q} f(a_{p+q}) = 0.$$

(*) 初学者可以只阅读本节的这一小节, 后面几小节将不会用到. 不过第 2 小节到第 5 小节的阅读即使对于初学者也是极其有益的练习.

而值 $f(a_{p+1}) = b_1, \cdots, f(a_{p+q}) = b_q$ 根据假设是线性无关的, 故得

$$\lambda_{p+1} = 0, \cdots, \lambda_{p+q} = 0.$$

原来的关系缩减为

$$\lambda_1 a_1 + \cdots + \lambda_p a_p = 0,$$

由于 $a_1, \cdots, a_p$ 是线性无关的, 因此有

$$\lambda_1 = 0, \cdots, \lambda_p = 0,$$

这就结束了证明.

注 1 后面将会看到 (§19, 定理 13), 当 K 是一个域时, 就可以谈论 K 上的有限维向量空间的 “维数”, 定理 1 意味着



$$\dim(\mathrm{L}) = \dim[\mathrm{Ker}(f)] + \dim[\mathrm{Im}(f)].$$

不言而喻, 在定理 1 的证明中假定 K 是域或交换域是不起任何作用的, 反之, 本节后面表明, 如果想开发这个定理的所有推论, 最一般的陈述则完全是本质的.

2. Noether 环上的有限生成模

设 I 是环 K 的一个*左理想* (即看作左 K-模的 K 的一个子模). 我们曾经说过 (§11, 例 6) I 是*有限生成的*, 如果它作为左 K-模是有限生成的, 即存在 I 的有限个元素 $x_1, \cdots, x_n$, 使得 I 是能够表示成形式 $u_1x_1 + \cdots + u_nx_n$ 的 K 的元素的集合.

我们说一个环 K 是**左 Noether 环**, 如果 K 的所有左理想是有限生成的. 同样通过考虑右理想定义右 Noether 环. 当 K 是交换环时 (这类问题中的最重要情形), 就简单地说 K 是 **Noether 环**.

例 1 一个域 K 是一个 Noether 环 (因为它的仅有的理想是 {0} 和 K, 它们显然是有限生成的). 一个主理想整环 (§8, 例 10) 也是 Noether 环.

主理想整环之外, Noether 环的最重要的例子是系数在一个域内的 n 个未知元的多项式环 (参见 §32, 习题 27). 这些环 —— 当 $n \geqslant 2$ 时它们不是主理想整环 —— 在 “代数流形” 即用代数方程定义的 “曲线”、“曲面” 等的研究中起着基本的作用.

定理 2 *设 K 是一个环, 下列性质是等价的:*

a) 环 K 是左 Noether 环.

b) 所有有限生成左 K-模的所有子模本身是有限生成的.

直接看出 b) 蕴含 a): 事实上, 左 K-模 K 是有限生成的, 因此, 如果 b) 满足, 它的子模 (即 K 的左理想) 一定是有限生成的.

现在证明 a) 蕴含 b). 分两步进行: 首先证明对于所有整数 $n \geqslant 1$, K^n 的所有子模是有限生成的; 然后由此推导出一般情形的 b).

我们证明 K^n 的子模 L 是有限生成的. 如果 $n=1$, 这个断言不是别的, 正是假设 a). 假定要证明的性质对于 $n-1$ 成立, 我们要对于 n 进行归纳推理. 为此, 通过 $f(\xi_1,\cdots,\xi_n)=\xi_n$ 定义一个同态 $f:\mathrm{L}\to\mathrm{K}$, 这是从 L 到 K 内的一个同态. 为了证明 L 是有限生成的, 只需 (定理 1) 指出 $\mathrm{Im}(f)$ 和 $\mathrm{Ker}(f)$ 是有限生成的. 而 $\mathrm{Im}(f)$ 是 K 的一个子模, 根据假设 a) 是有限生成的. 至于 $\mathrm{Ker}(f)$, 这是 K^n 的由关系 $\xi_n=0$ 定义的子模的一个子模, 而 K^n 的这个子模, 显然同构于 K^{n-1}, 根据归纳假设它的所有子模, 其中包括 $\mathrm{Ker}(f)$, 都是有限生成的.

借助 a) 我们已经证明, K^n 的所有子模是有限生成的. 取一个左 K-模 M 和它的一个子模 M′. 我们要指出, 如果 M 是有限生成的, 则 M′ 也是有限生成的. 由于 M 是有限生成的, 存在 (§12, 定理 3 的推论 2) 一个整数 n 和从 K^n 到 M 上的一个同态 f. 考虑 $\mathrm{L}=f^{-1}(\mathrm{M}')$, 由于 f 是满射的, f 诱导出从 L 到 M′ 上的一个同态. 根据刚才已经证明的结论, L 是有限生成的, 从而 M′ 是有限生成的 (更精确地说, 如果 L 由向量 $a_i, 1\leqslant i\leqslant p$ 生成, 那么 M′ 由 $f(a_i), 1\leqslant i\leqslant p$ 生成). 这就完成了证明.

3. 主理想整环上的自由模的子模

证明定理 2 所使用的方法还引导出下列结果:

定理 3 设 K 是一个环, 下列性质是等价的:

a) 对于 K 的所有非零左理想 I, 存在一个 $a\in\mathrm{I}$, 使得从 K 到 I 内的映射 $x\to xa$ 是双射.

b) 如果一个左 K-模 M 具有由 n 个向量组成的一个基, 并且 M′ 是 M 的一个子模, 则存在一个整数 $p\leqslant n$, 使得 M′ 具有由 p 个向量组成的一个基.

为了证明 b) 蕴含 a), 取 M = K, 那么 M 具有由一个向量组成的一个基; 从而 M 的所有子模 (即 K 的所有左理想) 具有一个由零个或一个向量组成的一个基, 这正是所陈述的 a).

现在证明 a) 蕴含 b). 显然 b) 等价于下列断言:

b′) 对于所有整数 $n\geqslant 1$ 和 K^n 的所有子模 L, 存在一个整数 $p\leqslant n$, 使得 L 同构于 K^p.

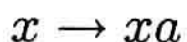
对于 $n=1$, 显然 b′) 归结为假设 a), 于是假定 b′) 对于 $n-1$ 成立而采用对 n 的归纳推理. 为此设 L 是 K^n 的一个子模, 并且跟前一小节一样, 由 $f(\xi_1,\cdots,\xi_n)=\xi_n$ 定义一个同态 $f:\mathrm{L}\to\mathrm{K}$. f 的像是 K 的一个子模, 因而对于一个整数 $r\leqslant 1$ 同构于 K^r. f 的核同构于 K^{n-1} 的一个子模, 故根据归纳假设, 对于一个整数 $s\leqslant n-1$, 它同构于 K^s. 利用定理 1, 我们推导出 L 同构于 K^{r+s}, 由于

$$r+s\leqslant 1+(n-1)=n,$$

证明完成.



注 2 当 K 是一个主理想整环时条件 a) 成立, 所谓主理想整环, 是一个交换的整环, 其所有理想都是主理想. 事实上, 设 I 是 K 的一个理想, 对于一个 $a \in \mathrm{K}$ 我们有 $\mathrm{I} = \mathrm{K}a$, 因此从 K 到 I 内的映射



$$x \to xa$$

是满射的. 再者, 其核由使得 $xa = 0$ 的 x 组成, 但如果 I 不缩减为零 (此时显然有 $a \neq 0$), 由于 K 是整环, 必有 $x = 0$. 故所考虑的映射是单射, 这就证明假设 a) 是满足的.

不言而喻, 当 K 是一个域 (交换或否) 时假设 a) 也满足, 因为这时 K 的仅有的非零左理想是 K, 只需在假设 a) 的陈述中取 $a = 1$.

在实际中, 这是仅有的使用定理 3 的两个情形.

注意到环 $\mathbf{Z}$ 是主理想整环, 定理 3 蕴含下列结果:

推论 $\mathbf{Z}^n$ 的所有子群对于一个整数 $p \leqslant n$ 同构于 $\mathbf{Z}^p$.

当然, 当用域代替 $\mathbf{Z}$ 时有类似的结果, 不过关于这种情形在下一节有更精细的结果.

4. 应用到线性方程组

设 K 是一个环, 而 f 是从右 K-模 K^n 到右 K-模 K^p 的一个同态. 设

$$(\alpha_{ij})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p}$$

是 f (关于 K^n 和 K^p 的典范基) 的矩阵. 如果 $x = (\xi_1, \cdots, \xi_n)$ 是 K^n 的一个向量, 它的像

$$y = f(x) = (\eta_1, \cdots, \eta_p)$$

由下列关系给定:

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{1n}\xi_n = \eta_1, \\ \cdots\cdots\cdots\cdots \\ \alpha_{p1}\xi_1 + \cdots + \alpha_{pn}\xi_n = \eta_p. \end{cases}$$

由此可见 f 的核是满足关系

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{1n}\xi_n = 0, \\ \cdots\cdots\cdots\cdots \\ \alpha_{p1}\xi_1 + \cdots + \alpha_{pn}\xi_n = 0 \end{cases} \tag{1}$$

的向量 $(\xi_1, \cdots, \xi_n)$ 组成的 K^n 的子模. 我们说 (1) 是系数在 K 内的 n 个未知元 p 个齐次线性方程的方程组.

设 L 是由 (1) 的解组成的 K^n 的子模. 我们要用线性方程的语言解释 L 的下列性质:

(i) L 是有限生成的.

(ii) 存在一个整数 $r \leqslant n$, 使得 L 同构于 K^r.

只要 K 是右 Noether 环 (定理 2), 第一个性质成立; 只要 K 是一个域或一个主理想整环 (定理 3), 第二个性质成立.

假定 L 是有限生成的; 那么存在 L 的有限数目的向量

$$\begin{cases} x^1 = (\xi_1^1, \cdots \xi_n^1), \\ x^2 = (\xi_1^2, \cdots \xi_n^2), \\ \cdots\cdots\cdots\cdots \\ x^r = (\xi_1^r, \cdots \xi_n^r), \end{cases} \tag{2}$$

使得 L 的元素

$$x = (\xi_1, \cdots, \xi_n)$$

可以写成下列形式

$$x = x^1\tau_1 + \cdots + x^r\tau_r, \tag{3}$$

其中 $\tau_1, \cdots, \tau_r$ 是任意标量. 考虑到 (2), 关系 (3) 显然等价于以下关系组:

$$\begin{cases} \xi_1 = \xi_1^1\tau_1 + \cdots + \xi_1^r\tau_r, \\ \qquad \cdots\cdots\cdots\cdots \\ \xi_n = \xi_n^1\tau_1 + \cdots + \xi_n^r\tau_r. \end{cases} \tag{4}$$

如此看来, 当 K 是一个右 Noether 环时, 存在常值 $\xi_i^j \in K$, 使得方程组 (1) 的解是标量组 $\xi_1, \cdots, \xi_n$, 它们可以写成形式 (4), 其中的 $\tau_j \in K$ 是任意标量. 我们把这个结果表述为: 组 (1) 的解依赖于有限个任意参量 (即出现在 (4) 中的 $\tau_1, \cdots, \tau_r$).

当 K 是一个域或一个主理想整环时, 甚至可以假定向量 (2) 组成 L 的一个基, 那么每个 $x \in L$, 按唯一的方式写成形式 (3). 即如果 K 是一个域或一个主理想整环, 则存在常量

$$\xi_i^j \in K \quad (1 \leqslant i \leqslant n, 1 \leqslant j \leqslant r \leqslant n),$$

使得由公式 (4) 给出的映射 $(\tau_1, \cdots, \tau_r) \to (\xi_1, \cdots, \xi_n)$ 是从 K^r 到 (1) 的解的集合上的一个双射.

当 K 是一个域时, 后面将给出更完备和更清晰的结果.

5. Noether 环的其他特征

在这一小节, 我们要给出 Noether 环的几个特征性质. 先引进下列术语.

设 (A_n) 是集合 X 的子集的一个序列, 说这是一个**递增序列**, 如果

$$A_n \subset A_{n+1} \quad \text{对于所有 } n;$$

说这是一个**稳定序列**, 如果存在一个整数 p, 使得

$$A_n = A_{n+1} \quad \text{对于 } n \geqslant p,$$

于是有 $A_p = A_{p+1} = A_{p+2} = \cdots$.

另外, 考虑 X 的子集的集合 F, 说一个 $A \in F$ 是 F 的**极大元素**, 如果

$$\text{关系 } A \subset B \text{ 并且 } B \in F \text{ 蕴含 } A = B,$$

换句话说, F 不含有严格大于 A 的集合 (这并不是说所有的 $B \in F$ 一定包含于 A 内).

定理 4 设 K 是一个环, 以下性质是等价的:
(AN1) K 是左 Noether 环 (即 K 的所有左理想是有限生成的);
(AN2) K 的左理想的所有递增序列是稳定的;
(AN3) K 的左理想的所有非空集合至少具有一个极大元素.

我们证明 (AN1) 蕴含 (AN2). 设 (I_n) 是 K 的左理想的一个递增序列, I_n 的并集 I 还是一个左理想 (§10, 定理 1). 由 Noether 环的定义, I 是由有限个元素 $x_1, \cdots, x_r$ 生成的, 根据并集的定义, 存在整数 $p_1, \cdots, p_r$, 使得 $x_1 \in I_{p_1}, \cdots, x_r \in I_{p_r}$. 令 $p = \max(p_1, \cdots, p_r)$, 那么 I_p 由于包含 $I_{p_1}, \cdots, I_{p_r}$ 而含有 $x_1, \cdots, x_r$, 故 I_p 包含由 $x_1, \cdots, x_r$ 生成的理想, 即 I. 对于 $n \geqslant p$ 有 $I_p \subset I_n \subset I \subset I_p$, 故 $I_p = I_n$, 这就证明了 (AN2).

现在证明 (AN2) 蕴含 (AN3). 设 F 是 K 的左理想的一个非空集合. 如果 F 没有任何极大元素, 那么对于所有 $I \in F$, 可以找到一个严格包含 I 的 J. 选定一个 $I_1 \in F$ 之后, 存在一个严格包含 I_1 的 I_2, 从而严格包含 I_2 的一个 I_3, 如此继续下去, 采用这种方式显然得到严格递增的 K 的左理想的一个序列, 这违背了性质 (AN2).

最后要证明 (AN3) 蕴含 (AN1). 设 I 是 K 的一个左理想, 设 F 是有限生成的包含于 I 内的 K 的左理想的集合, F 是非空的 (例如它含有 K 的理想 $\{0\}$). 根据 (AN3), 集合 F 具有一个极大元素 J. 设 $(x_1, \cdots, x_n)$ 是 J 的一组生成元. 对于所有 $x \in I$, 由 $x_1, \cdots, x_n$ 和 x 生成的左理想 I' 是有限生成的并且包含于 I 内, 从而 I' 属于 F 并且包含 J, 由于 J 是 F 的一个极大元素, I' 这个左理想只能是 J 本身. 于是对于所有 $x \in I$ 有 $x \in J$, 故得 $I = J$, 这就证明了 I 是有限生成的, 并且完成了定理的证明.

注 3 我们说环 K 的一个左 (对应的, 右、双侧) 理想 I 是**极大的**, 如果 $I \neq K$, 并且包含 I 的 K 的仅有的左 (对应的, 右、双侧) 理想是 I 和 K. 借助集合论的相当复杂的推理可以证明, 如果 K 是一个环, K 的所有异于 K 的左 (对



应的, 右、双侧) 理想 I 至少包含于 K 的一个极大左 (对应的, 右、双侧) 理想内 (Krull 定理).

现在这个结果在代数中起着基本的作用, 如果 K 是 Noether 环, 则可以用初等的方法证明: 为此只需利用定理 4 的断言 (AN3) 到 K 的左 (对应的, 右、双侧) 理想 J 的集合, 这里的 J 满足

$$\mathrm{I} \subset \mathrm{J}, \quad \mathrm{J} \neq \mathrm{K}.$$

§18 习题

¶1. 证明加法群 $\mathbf{Q}^n$ 的所有有限生成子群具有至多 n 个元素的一个基.

¶2. 存在含有 $\mathbf{Z}^n$ 的给定元素 $(a_1, \cdots, a_n)$ 的 $\mathbf{Z}^n$ 的一个基, 必须并且只需诸整数 a_i 是互素的 (选择 $u_i \in \mathbf{Z}$, 使得 $\sum u_i a_i = 1$, 并且考虑由方程 $\sum u_i x_i = 0$ 定义的 $\mathbf{Z}^n$ 的子群).

更一般的, 设 K 是一个环. 为了存在 K^n 的一个含有给定的 $a \in \mathrm{K}^n$ 的基, 必须存在 K^n 上的一个线性型 f, 使得 $f(a) = 1$, 并且如果 K 是主理想整环 (参见 §17, 习题 9), 这个条件还是充分的.

¶3. 存在一个其第一行为

$$a_{11}, a_{12}, \cdots, a_{1n}$$

的矩阵 $\mathrm{U} \in \mathrm{GL}(n, \mathbf{Z})$, 必须且只需整数 $a_{11}, a_{12}, \cdots, a_{1n}$ 是互素的.

¶4. 设 L 和 M 是交换的 Noether 环 K 上的两个有限生成模. 证明 K-模 $\mathrm{Hom}_{\mathrm{K}}(\mathrm{L}, \mathrm{M})$ 是有限生成的 (构造从这个模到 M 的一个适当的幂内的一个单射的同态).

¶5. 设 M 是 Noether 环 K 上的有限生成的一个模. 证明 M 的所有递增序列是稳定的, 并且 M 的所有子模的非空集合至少具有一个极大元素.

¶¶6. 设 $\mathfrak{a}$ 是交换的 Noether 环 K 的一个理想. 证明存在 K 的素理想的一个有限序列 $\mathfrak{p}_1, \cdots, \mathfrak{p}_n$ (不必是两两不同的), 使得

$$\mathfrak{p}_1, \cdots, \mathfrak{p}_n \subset \mathfrak{a}.$$

(进行归谬推理, 考虑不具有这个性质的理想的集合的极大元素, 并且对于它应用 §8 的习题 11.)

¶¶7. 设 K 是一个交换环, 而 A 是 K 的一个子环, 以 K 作为分式域 (即所有的 $x \in \mathrm{K}$ 是 A 的两个元素的商). 下面利用 §11 的习题 14 的定义和结果.

a) 如果 A 是一个 Dedekind 整环, 则 A 是 Noether 环.

b) 假定 A 是 Dedekind 整环, 并且环 A 的所有极大理想是可逆的. 证明 A 的所有理想是有限个极大理想的乘积 (类似于前一个小题的方法). 由此推出 A 是 Dedekind 整环.

c) 设 A 是 Dedekind 整环. 证明 A 的所有分式理想可以写成 A 的素理想的 (正的或负的) 幂的有限乘积的形式, 并且如果不计因子的次序这种分解是唯一的.

d) 设 $\mathfrak{p}$ 是 Dedekind 整环 A 的一个素理想. 对于所有非零的 $x \in \mathrm{K}$, 用 $v_{\mathfrak{p}}(x)$ 表示将 A 的分式理想 Ax 表示成素因子乘积的分解中 $\mathfrak{p}$ 的指数 (可能是零); 并且定义 $v_{\mathfrak{p}}(0) = \infty$. 证明函数 $v_{\mathfrak{p}}$ 是域 K 的离散赋值 (§8, 习题 6).

¶8. 设 K 是一个交换的 Noether 环, M 是一个有限生成的 K-模, 而 u 是 M 内的比例为 a 的位似变换, 其定义是

$$u(x) = ax \quad \text{对于所有的 } x \in \mathrm{M}.$$

a) 证明存在一个整数 $p \geqslant 0$, 使得对于所有 $n \geqslant p$ 有 $\mathrm{Ker}(u^n) = \mathrm{Ker}(u^{n+1})$ (利用习题 5).

b) 证明对于所有 $n \geqslant p$ 有 $\mathrm{Im}(u^n) \cap \mathrm{Ker}(u) = \{0\}$.

c) 称模 M 是**准素的**, 如果在 M 内, 所有位似或是单射的, 或是幂零的. 证明当 M 具有下列性质时它就是准素的: M 的两个非零的子模的交集从不是零.

d) 设 $\mathfrak{q}$ 是环 K 的一个理想. $\mathfrak{q}$ 是准素的 (§8, 习题 13), 必须并且只需将商 $\mathrm{K}/\mathfrak{q}$ 看作 K-模是准素的.

e) 环 K 的一个理想 $\mathfrak{q}$ 是**不可约的**, 如果 $\mathfrak{q} \neq \mathrm{K}$, 并且对于 K 的两个任意理想 $\mathfrak{a}$ 和 $\mathfrak{b}$,

$$\text{关系 } \mathfrak{a} \cap \mathfrak{b} = \mathfrak{q} \text{ 蕴含 } \mathfrak{a} = \mathfrak{q} \text{ 或 } \mathfrak{b} = \mathfrak{q}.$$

证明 Noether 环的所有不可约理想是准素的.

¶¶9. 设 K 是一个交换的 Noether 环.

a) 证明所有理想 $\mathfrak{a} \neq \mathrm{K}$ 是有限个不可约理想的交集 (习题 6 的方法).

b) 由此推出交换的 Noether 环的所有理想是有限个准素理想的交集 (Emmy Noether).

¶¶10. 设 M 是交换的 Noether 环 K 上的一个有限生成的模. 称 K 的一个素理想 $\mathfrak{p}$ **是相伴于 M 的**, 如果存在一个 $x \in \mathrm{M}$, 使得 $\mathfrak{p}$ 是 x 在 M 内的零化子 (即关系 $a \in \mathfrak{p}$ 等价于关系 $ax = 0$).

a) 如果 $\mathrm{M} \neq \{0\}$, 则至少存在一个相伴于 M 的素理想 (考虑 M 的非零元素的零化子, 并且从中选取一个极大的).

b) 存在 M 的子模的递增序列

$$0 = \mathrm{M}_0 \subset \mathrm{M}_1 \subset \cdots \subset \mathrm{M}_r = \mathrm{M}$$

和 K 的素理想 $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$, 使得对于满足条件 $1 \leqslant i \leqslant r$ 的所有 i, $\mathrm{M}_i/\mathrm{M}_{i-1}$ 同构于商模 $\mathrm{K}/\mathfrak{p}_i$ (注意如果一个模的一个元素 x 的零化子是 K 的一个理想 $\mathfrak{a}$, 那么由 x 生成的子模 $\mathrm{K}x$ 同构于 $\mathrm{K}/\mathfrak{a}$).

c) 沿用前一个问题的记号, 所有相伴于 M 的素理想是诸 $\mathfrak{p}_i$ 中的一个.

d) 设 $u(x) = ax$ $(a \in \mathrm{K})$ 是 M 内的位似. u 是单射, 必须并且只需 a 不属于相伴于 M 的任意素理想; u 是幂零的, 必须并且只需 a 属于相伴于 M 的所有素理想.

e) 下面取 $\mathrm{M} = \mathrm{K}$, 并且用 $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$ 表示相伴于 M 的素理想. 证明

$$\mathfrak{p}_1 \cup \cdots \cup \mathfrak{p}_r$$

是 K 内的零因子的集合, 并且

$$\mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r$$

是 K 的幂零元素的集合.

证明 K 的所有素理想含有诸 $\mathfrak{p}_i$ 中的一个. 由此推出在一个 Noether 环内, 所有素理想的交集是幂零元素的集合 (事实上这个结果可以推广到所有的交换环: §28, 习题 9).

¶¶11. 设 $\mathfrak{a}$ 是交换的 Noether 环 K 的一个理想, 假定 $\mathfrak{a} \neq \mathrm{K}$. K 的所有素理想 $\mathfrak{p}$, 如果它包含 $\mathfrak{a}$, 并且不包含其他任何包含 $\mathfrak{a}$ 的素理想, 则称 $\mathfrak{p}$ 为 $\mathfrak{a}$ 的**极小素理想**. 证明 $\mathfrak{a}$ 的极小素理想的集合是有限的, 并且它们的交集是 $\mathfrak{a}$ 的根 (§8, 习题 12) (利用前一个习题的问题 e) 到 $\mathrm{K}/\mathfrak{a}$).

(Noether 环由 Emmy Noether 在 1920 年发现, 并且是现代 "抽象" 代数的主要出发点之一. 它们蓬勃发展的理论是代数几何的基础, 而且在其他方面也有应用, 尤其是在多元复变量的解

析函数理论中. 事实上, 这些环的发现无疑是近代最有用的数学发现之一. 在 Noether 环引进之前, 人们局限于证明多项式环的一些性质, 这就经常导致比现今所知道的证明复杂得多的证明, 这是由于人们还没有洞察到能够简化证明的基本思想, 即本节的定理 2 和 4.)

§19　维数概念

1. 基的存在性

下面的结果已经在 §11 第 4 小节陈述过:

定理 1　域上的所有有限维(*)向量空间具有一个基.

这个定理显然是下述更精确的结果的一个推论:

定理 2　设 M 是一个域上的有限维向量空间, X 是 M 的生成元的有限集合, 而 A 是 X 的一个子集, 假定 A 的元素是线性无关的, 那么存在 M 的一个基 B, 使得

$$A \subset B \subset X.$$

事实上, 考虑 X 的所有这样的子集, 这些子集包含 A 并且是自由的 (即线性无关的); 这样的子集是存在的, 首先 A 自己就是其中一个. 在这些子集中, 我们考虑元素数目最多的那些子集, 设 B 是其中的一个. 为了证明定理 2, 只需证明 B 是 M 的一个基, 或同样的, B 生成 M, 因为由 B 的选择, 它是自由的.

由于 X 生成 M, 为了证明 B 生成 M, 只需证明所有 $x \in X$ 是 B 的元素的线性组合. 如果 $x \in B$, 这是显然的, 为了证明上述断言, 我们假定 $x \notin B$.

集合 $B' = B \cup \{x\}$ 包含于 X 内, 并且比 B 确实含有更多的元素. 还有 $A \subset B'$, 因此, 我们看到 B′ 不可能是自由的. 于是如果用 $x_1, \cdots, x_r$ 表示 B 的不同元素, 将存在关系

$$\lambda_1 x_1 + \cdots + \lambda_r x_r + \lambda x = 0, \tag{1}$$

其中的 $\lambda_1, \cdots, \lambda_r, \lambda$ 不全为零. 甚至有 $\lambda \neq 0$, 因为如果不然, (1) 将缩减为元素 $x_i \in B$ 之间的一个非平凡的线性关系, 这与 B 是自由的相矛盾.

由于 λ 不是零, 又由于基础环 K 是一个域, λ 在 K 内是可逆的, 用 λ 的逆元乘 (1) 左端即得

$$x = -\lambda^{-1}\lambda_1 x_1 - \cdots - \lambda^{-1}\lambda_r x_r.$$

这样我们就证明了所有 $x \in X$ 是 B 的元素的线性组合, 由此即得定理.

定理 1 和定理 2 (在 §11 之后就可以证明它们) 有重要的推论, 除下面两个推论外, 在第 2 小节和第 3 小节还有几个.

(*) 本小节的结果事实上对于所有向量空间, 即使是无限维的都是有效的, 但是一般情形的证明在这里复述实在太困难了.

推论 1 设 M 是一个域上的有限维向量空间. M 的一个给定集合成为 M 的基的一个子集, 必须并且只需它们是线性无关的.

如果存在 M 的一个基含有给定向量 $x_1, \cdots, x_p$, 它们显然是线性无关的. 反之, 如果这个条件满足, 为了得到含有给定向量的 M 的基, 选择 M 的生成元的一个有限组 G, 并且应用定理 2 到集合

$$\mathrm{X} = \mathrm{G} \cup \{x_1, \cdots, x_p\}, \quad \mathrm{A} = \{x_1, \cdots, x_p\}.$$

推论 2 设 M 是一个域上的有限维向量空间. M 的所有向量子空间是 M 内的一个直和项.

事实上, 设 M′ 是 M 的子空间. 根据 §18 定理 2, M′ 是有限维的, 于是 (根据定理 1) 具有一个基 $\{x_i\}_{1\leqslant i\leqslant p}$. 应用推论 1 到这些向量和 M, 我们发现存在向量 $x_{p+1}, \cdots, x_r$, 使得 $x_i (1 \leqslant i \leqslant r)$ 组成 M 的一个基. 那么显然由 $x_{p+1}, \cdots, x_r$ 生成的 M 的子空间就是 M′ 在 M 内的补空间.



注 1 上面推论 2 的证明利用了域 K 上的有限维空间 M 的一个子空间就是域 K 上的有限维空间这个事实, 而这个事实从一个域是 Noether 环 (这时利用 §18 定理 2) 或主理想整环 (这时利用 §18 定理 3 进行推理) 得到. 但是显然可以给这个事实一个直接且初等的证明 (它本质上与 §18 定理 2 和 3 的证明没有区别), 以下便是这样证明.

由于 M 拥有一个基 (定理 1), 可以假定 $\mathrm{M} = \mathrm{K}^n$. 如果 $n = 1$, M 的仅有的子空间是 $\{0\}$ 和 M, 由于 K 是一个域, 它们是有限维的. 在一般情形, 设 M′ 是 K^n 的一个子空间, 并且把 K^{n-1} 与 K^n 的使得 $\xi_n = 0$ 的向量 $(\xi_1, \cdots, \xi_n)$ 组成的子空间等同. 如果 $\mathrm{M}' \subset \mathrm{K}^{n-1}$, 可以假定 M′ 是有限维的 (关于 n 进行归纳推理). 如果 M′ 不包含于 K^{n-1} 内, 在 M′ 内选择一个向量

$$\alpha = (\alpha_1, \cdots, \alpha_n), \quad \text{其中 } \alpha_n \neq 0,$$

并且令 $\mathrm{M}'' = \mathrm{M}' \cap \mathrm{K}^{n-1}$. 对于 M′ 的所有元素

$$x = (\xi_1, \cdots, \xi_n),$$

由于 α_n 在 K 内是可逆的, 可以写出

$$x = \xi_n \alpha_n^{-1} a + y,$$

其中向量 y 的最后一个分量为零, 从而属于 K^{n-1}, 并且显然属于 M′, 于是属于 M″. 我们立刻发现 M′ 是由 a 生成的子空间 $\mathrm{K}a$ 与 M″ 的直和, 而 M″ 是 K^{n-1} 的一个子空间, 根据归纳假设是有限维的, 于是 M′ 也是有限维的, 这就完成了证明.

在文献中还可以发现许多其他的证明方法. 例如, 可以首先证明定理 2 和 6, 然后证明定理 10, 11 和 12. 本注所证明的结果立刻由定理 12 推演出来.

2. 由线性方程组定义向量子空间

设 L 是一个域 K 上的有限维向量空间, M 是 L 的一个子向量空间. 在 L 的对偶空间 L^* 内, 考虑 L 上的满足关系

$$f(x)=0 \quad \text{对于所有 } x\in \mathrm{M}$$

的线性型 f 的集合, 把它记作

$$\mathrm{M}^0.$$

这个集合是 L^* 的向量子空间, 因为如果它含有两个线性型 f 和 g, 并且如果令 $h=\alpha f+\beta g$, 其中 α 和 β 是任意标量, 则有

$$h(x)=\alpha\cdot f(x)+\beta\cdot g(x)=0 \quad \text{对于所有 } x\in \mathrm{M},$$

故 $h\in \mathrm{M}^0$, 这就证明了我们的断言.

我们说 M^0 是 M **在 L^* 内的零化子**. 以下结果表明可以从子空间 M^0 重建子空间 M.

定理 3 设 L 是一个有限维向量空间, M 是 L 的一个向量子空间, 而 M^0 是 M 在 L^* 内的零化子. $x\in \mathrm{L}$ 是在 M 内, 必须且只需对于所有线性型 $f\in \mathrm{M}^0$ 有

$$f(x)=0. \tag{2}$$

条件的必要性是平凡的, 我们要证明它是充分的. 正如在定理 2 的推论 2 的证明中看到的, 存在 L 的一个基 $(a_i)_{1\leqslant i\leqslant r}$ 和一个整数 $p\leqslant r$, 使得 M 是由 $a_1,\cdots,a_p$ 生成的. 设 $f_1,\cdots,f_r$ 是关于基 $a_1,\cdots,a_r$ 的坐标函数, 显然 M 由关系

$$f_{p+1}(x)=\cdots=f_r(x)=0 \tag{3}$$

定义, 换句话说, M^0 含有线性型 $f_{p+1}(x),\cdots,f_r(x)$, 并且关系 (3) 刻画了 M 的元素的特征. 如果条件 (2) 对于所有 $f\in \mathrm{M}^0$ 成立, 显然条件 (3) 因此成立. 定理证毕.



注 2 其实, 上面构造的线性型 $f_{p+1}(x),\cdots,f_r(x)$ 组成 M^0 的一个基. 事实上, 由于它们是线性无关的 (线性型 $f_1,\cdots,f_r$ 组成与 L 的基 $a_1,\cdots,a_r$ 对偶的 L^* 的基), 只需证明它们生成 M^0. 考虑一个线性型 $f\in L^*$, 令 $f(a_i)=\alpha_i$, 则对于

$$x=a_1\xi_1+\cdots+a_r\xi_r,$$

有

$$f(x)=f(a_1\xi_1+\cdots+a_r\xi_r)=\alpha_1\xi_1+\cdots+\alpha_r\xi_r,$$

由于 $\xi_i = f_i(x)$, 我们得到

$$f = \alpha_1 f_1 + \cdots + \alpha_r f_r, \quad 其中\ \alpha_i = f(a_i).$$

如果 $f \in \mathrm{M}^0$, 则有 $f(a_1) = \cdots = f(a_p) = 0$, 故

$$f = \alpha_{p+1} f_1 + \cdots + \alpha_r f_r,$$

这就证明了 $f_{p+1}, \cdots, f_r$ 生成 M^0.

反之, 如果线性型 $f_i(1 \leqslant i \leqslant m)$ 生成 M^0, 那么 M 的元素由关系

$$f_1(x) = \cdots = f_m(x) = 0$$

刻画其特征. 事实上, 所有 $f \in \mathrm{M}^0$ 可以表示成形式 $f = \alpha_1 f_1 + \cdots + \alpha_m f_m$, 其系数 $\alpha_i \in \mathrm{K}$, 那么显然所考虑的方程蕴含 $f(x) = 0$, 根据定理 3 即得 $x \in \mathrm{M}$.

定理 3 的主要用途在于导出我们刚证明的性质, 值得给它一个明晰的陈述:

推论 1 设 $(f_i)_{1\leqslant i\leqslant m}$ 是 M^0 的一个生成元组, 那么 M 的元素由方程

$$f_1(x) = \cdots = f_m(x) = 0$$

刻画其特征.

这里还有定理 3 的一个重要推论:

推论 2 设 L 是一个有限维向量空间, 而 M 是 L 的一个向量子空间. 有 $\mathrm{M} \neq \mathrm{L}$, 必须且只需存在 L 上的一个非零线性型, 使得

$$f(x) = 0 \quad 对于所有\ x \in \mathrm{M}.$$

如果 $\mathrm{M} = \mathrm{L}$, 那么显然 $\mathrm{M}^0 = \{0\}$. 反之, 如果 $\mathrm{M}^0 = \{0\}$, 那么定理 3 表明 $\mathrm{M} = \mathrm{L}$. 于是 $\mathrm{M} = \mathrm{L}$ 和 $\mathrm{M}^0 = \{0\}$ 是等价的, 从而 $\mathrm{M} \neq \mathrm{L}$ 等价于 $\mathrm{M}^0 \neq \{0\}$, 这就证明了推论.

3. 线性方程组相容性条件

线性方程组理论将在 §20 仔细研究, 但是现在就可以解决判断在什么条件下一个线性方程组具有解的问题. 这将是本小节要证明的两个定理所要面对的问题.

定理 4 设 $f_1, \cdots, f_r$ 是一个域 K 上的向量空间 M 上的线性型. 下列条件是等价的:

a) $f_1, \cdots, f_r$ 是线性无关的.

b) 对于任意标量 $\beta_1, \cdots, \beta_r \in \mathrm{K}$, 至少存在一个 $x \in \mathrm{M}$, 满足

$$\begin{cases} f_1(x) = \beta_1, \\ \cdots\cdots\cdots\cdots \\ f_r(x) = \beta_r. \end{cases} \tag{4}$$

事实上, 考虑由

$$f(x) = (f_1(x), \cdots, f_r(x))$$

给定的同态 $f: \mathrm{M} \to \mathrm{K}^r$. 条件 b) 表示 f 是满射的. 而 $f(\mathrm{M})$ 是 K^r 的一个线性子空间, 根据定理 3 的推论 2, 一切归结为如果 K^r 上的一个线性型 u 在子空间 $f(\mathrm{M})$ 上是零, 则 $u = 0$.

如果

$$u(\xi_1, \cdots, \xi_r) = \lambda_1\xi_1 + \cdots + \lambda_r\xi_r$$

是这样的一个线性型, 则对于所有 $x \in \mathrm{M}$ 有

$$u(f(x)) = \lambda_1 f_1(x) + \cdots + \lambda_r f_r(x);$$

说 u 在 $f(\mathrm{M})$ 上是零即表明线性型 f_i 满足线性关系

$$\lambda_1 f_1 + \cdots + \lambda_r f_r = 0.$$

如此看来, 所陈述的性质 a) 表明在 $f(\mathrm{M})$ 上为零的仅有的线性型 u 是 $u = 0$, 定理 4 证明完毕.

现在考虑线性方程组

$$\begin{cases} f_1(x) = \beta_1, \\ \cdots\cdots\cdots\cdots \\ f_n(x) = \beta_n, \end{cases} \tag{5}$$

其中的向量空间 M 上的线性型 $f_1, \cdots, f_n$ 不再必须是线性无关的. 在 M 的对偶空间 M^* 里, 这些线性型生成一个有限维向量空间 F, 根据定理 2, 可以从 $(f_i)_{1 \leqslant i \leqslant n}$ 抽出 F 的一个基. 如果有必要, 调整一下 f_i 的编号, 可以假定 $f_1, \cdots, f_r$ 组成 F 的一个基, 显然我们处于应用以下结果的条件下:

定理 5 设 $f_1, \cdots, f_n$ 是一个向量空间 M 上的线性型. 假定 $f_1, \cdots, f_r$ 是线性无关的, 并且有关系

$$f_j = \rho_{j1} f_1 + \cdots + \rho_{jr} f_r \quad \text{对于 } r+1 \leqslant j \leqslant n, \tag{6}$$

其中的系数 $\rho_{jk} \in \mathrm{K}$. 那么方程组 (5) 至少具有一个解 $x \in \mathrm{M}$, 必须且只需有

$$\beta_j = \rho_{j1}\beta_1 + \cdots + \rho_{jr}\beta_r \quad \text{对于 } r+1 \leqslant j \leqslant n. \tag{7}$$

条件 (7) 满足时, 方程组 (5) 与方程组

$$\begin{cases} f_1(x) = \beta_1, \\ \cdots\cdots\cdots\cdots \\ f_r(x) = \beta_r \end{cases} \tag{8}$$

具有同样的解.

关系 (6) 的意思是对于所有 $x \in \mathrm{M}$ 有

$$f_j(x) = \rho_{j1} f_1(x) + \cdots + \rho_{jr} f_r(x), \tag{6'}$$

于是显然, 如果存在一个 x 使得关系 (5) 成立, 关系 (7) 必然成立.

反之, 假定关系 (7) 成立, 考虑 (8) 的一个解 x, 那么按照 (6′),

$$f_j(x) = \rho_{j1}\beta_1 + \cdots + \rho_{jr}\beta_r,$$

考虑到 (7) 我们发现方程组 (5) 和 (8) 有同样的解. 由于根据前面的定理, (8) 实际上有解, 因而定理 5 证明完毕.

注 3 关系 (7) 对于方程组 (5) 至少具有一个解是必要且充分的, 称为方程组 (5) 的**相容性条件**. 定理 5 表明, 为了求线性方程组的解, 总可以归结为 (如果相容性条件满足) 其左端是线性无关的方程组. 这个结果自然应当补充以定理 4, 即一个线性无关的方程组总有解.



我们注意到在前面的陈述中没有假定 M 是有限维的; 这个假设在当前的行文中是多余的, 仅在证明 §20 的更精细的结果时才会用到.

4. 线性关系的存在性

我们现在要证明线性代数的基本结果.

定理 6 设 M 是域 K 上的有限维向量空间, 则 M 的所有基有同样数目的元素.

设 $(a_i)_{1\leqslant i\leqslant p}$ 和 $(b_j)_{1\leqslant j\leqslant q}$ 是 M 的两个基; 为了证明 $p = q$, 只需指出 $p \leqslant q$ 和 $q \leqslant p$. 由对称性理由, 甚至只需指出 $p \leqslant q$. 这显然会由下列陈述推导出来:

定理 7 设 M 是域 K 上的有限维向量空间, 而 p 是一个整数, 使得 M 具有 p 个向量组成的基. 为了 M 的 q 个向量是线性无关的, 必须 $q \leqslant p$.

不言而喻, 定理 7 的逆命题是错误的.

设 $(a_i)_{1\leqslant i\leqslant p}$ 是 M 的一个基, 而 $b_1, \cdots, b_q$ 是 M 的元素. 我们要指出, 一旦 $q > p$, 将存在向量 b_j 之间的一个非平凡的线性关系, 这就保证了定理 7 成立.

如果 $p = 0$, 要证明的断言是平凡的: 事实上, 这时我们有 $\mathrm{M} = \{0\}$, 而由于 $q \geqslant 1$, 显然 b_j 比如说满足关系

$$1 \cdot b_1 + 0 \cdot b_2 + \cdots + 0 \cdot b_q = 0.$$

现在假定定理对于 $p-1$ 成立, 我们要推出对于 p 它也成立. 设 M′ 是由 $a_1, \cdots, a_{p-1}$ 生成的 M 的子空间, 我们有关系

$$b_j = b'_j + \alpha_j a_p \quad (1 \leqslant j \leqslant q), \tag{9}$$

其中的 $b'_j \in \mathrm{M}'$, 并且标量 $\alpha_j \in \mathrm{K}$. 如果所有的 α_j 是零, 则 b_j 在 M' 内. 由于 M' 具有一个由 $p-1$ 个向量组成的基, 归纳假设用于 M', 而 $q>p$, 更加有 $q>p-1$, 我们发现在这种情形直接得到 b_j 之间的一个非平凡的线性关系的存在性.

余下要考察 α_j 不全为零的情形. 比如

$$\alpha_q \neq 0,$$

由于 K 是一个域, α_q 是可逆的. 那么 (9) 的最后一个关系给出

$$a_p = \alpha_q^{-1}(b_q - b'_q),$$

代入到 (9) 中的其他关系, 经过简单的计算即得

$$b_j - \nu_j b_q = b'_j - \nu_j b'_q \quad (1 \leqslant j \leqslant q-1), \tag{10}$$

其中的 $\nu_j = \alpha_j \alpha_q^{-1}$.

关系 (10) 指出 $q-1$ 个向量 $b_j - \nu_j b_q$ 在子空间 M' 内. 由于 $q>p$, 自然有 $q-1>p-1$, 归纳法假设指出存在线性关系

$$\lambda_1(b_1 - \nu_1 b_q) + \cdots + \lambda_{q-1}(b_{q-1} - \nu_{q-1} b_q) = 0,$$

其中的 $\lambda_1, \cdots, \lambda_{q-1}$ 不全为零. 而这个关系可以改写为

$$\lambda_1 b_1 + \cdots + \lambda_q b_q = 0, \quad 其中 \quad \lambda_q = -(\lambda_1 \nu_1 + \cdots + \lambda_{q-1} \nu_{q-1}),$$

$b_1, \cdots, b_q$ 之间的非平凡线性关系的存在性得以证明.

注 4 事实上, 定理 7 不仅对于域 (交换或否) 上的向量空间有效, 对于交换环上的模也是有效的. 借助行列式理论可以证明这一结果 (§23, 定理 5 的推论).

推论 设

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{1p}\xi_p = 0, \\ \qquad \cdots\cdots\cdots\cdots \\ \alpha_{n1}\xi_1 + \cdots + \alpha_{np}\xi_p = 0 \end{cases} \tag{11}$$

是系数在一个域 K 内 p 个未知元 $\xi_1, \cdots, \xi_p$ 的 n 个齐次线性方程的方程组. 如果 $p>n$ (即未知元的数目超过方程的数目), 则 (11) 至少具有一个非平凡解 (即未知元 ξ_i 不全为零的解).

考虑 K^n 的 p 个元素

$$a_i = (a_{1i}, \cdots, a_{ni}),$$

关系 (11) 显然等价于关系

$$a_1\xi_1 + \cdots + a_n\xi_n = 0.$$

因为方程组 (11) 有非零解, 因此充分和必要的是向量 a_i 是线性无关的. 而由定理 7, 当 $p > n$, 就是这种情形.

注 5 这个推论的直观解释如下: 由于方程组 (11) 总具有平凡解



$$\xi_1 = \cdots = \xi_p = 0,$$

所有的问题归结为发现保证 (11) 至少有两个解的条件, 即关系 (11) 不完全确定未知元 ξ_i 的条件. 如果我们对于 p 个未知元 ξ_i 施加至少 p 个条件, 那么就有极大的机会完全确定它们, 反之, 借助少于 p 个的条件确定它们, 则只有少许的可能性.

5. 维数概念

设 M 是域 K 上的有限维向量空间. 根据定理 6, 存在一个完全确定的整数 n, 使得 M 的所有的基都具有 n 个元素. 我们说 n 是**域 K 上的 M 的维数** (或仅当不可能有关于基础域 K 的任何混淆时简称为 **M 的维数**), 记作

$$\dim_{\mathrm{K}}(\mathrm{M}),$$

或当不可能有关于所选择的域的任何混淆时简单地记作 $\dim(\mathrm{M})$. 当 $\mathrm{M} = \{0\}$ 时取 $\dim(\mathrm{M}) = 0$ 是合适的.

注 6 一个复向量空间 M 可以看作一个实向量空间, 于是有关系



$$\dim_{\mathbf{R}}(\mathrm{M}) = 2 \cdot \dim_{\mathbf{C}}(\mathrm{M})$$

(习题 15), 这表明谈论维数时明确所选择的基础域是必须的.

注 7 还可以对于无限维向量空间定义维数概念, 不过要使用基数理论 (§5). 当 M 不是有限维时, 令 $\dim(\mathrm{M}) = +\infty$ 的初等方法看不出任何益处, 因为使用维数概念的这个过分简单的定义, 我们所面对的任何命题 (首先是下面的定理 8) 都不是真的, 而一个定义的益处, 在于它应导出一些有意义的定理(*), 当然还不止于此.

维数概念的主要益处集中体现在下列定理中:

定理 8 *设 L 和 M 是域 K 上的有限维向量空间. L 和 M 是同构的, 必须并且只需 $\dim(\mathrm{L}) = \dim(\mathrm{M})$.*

(*) 有时候在日常生活中也是这样. 比如, 如果我们的目的是证明 “Poldèvie 军队没有开枪射击 Poldèvie 人”, 只需引进下列定义: 在 Poldèvie, 称 Poldèvie 军队没有向其开枪射击的人为 Poldèvie 人. 这个过程, 虽然有一定的逻辑性, 但是表现出严重的不足: 人们又返回到 Poldèvie 军队的定义本身.

如果存在从 L 到 M 上的同构 f, f 把 L 的一个基映射到 M 的一个基上, 由此得到 $\dim(\mathrm{L}) = \dim(\mathrm{M})$. 反之如果这个条件满足, 设 n 是 L 和 M 的公共维数, 那么 (根据 §12 定理 3 的推论 1) L 和 M 都同构于 K^n, 于是它们彼此同构.

注 8　两个无限维向量空间总是同构是不正确的.

例 1　显然对于任意整数 n 有

$$\dim(\mathrm{K}^n) = n.$$

例 2　取 $\mathrm{K} = \mathbf{R}$, 取在通常空间内的起点为给定的点 O 的所有向量组成的向量空间为 M (§10, 例 2), 那么我们有 $\dim(\mathrm{M}) = 3$. 如果取位于一个平面上的过一个给定点 O 的向量的集合为 M, 则 $\dim(\mathrm{M}) = 2$. 最后取一条直线上的过一个给定点 O 的向量的集合为 M, 则 $\dim(\mathrm{M}) = 1$.

设 L 和 M 是域 K 上的有限维向量空间, 则有

$$\dim(\mathrm{L} \times \mathrm{M}) = \dim(\mathrm{L}) + \dim(\mathrm{M});$$

更精确地说, 如果 $(a_i)_{1 \leqslant i \leqslant p}$ 和 $(b_j)_{1 \leqslant j \leqslant q}$ 分别是 L 和 M 的基, 则 $p + q$ 个向量

$$(a_1, 0), \cdots, (a_p, 0), (0, b_1), \cdots, (0, b_q)$$

组成 $\mathrm{L} \times \mathrm{M}$ 的一个基.

例 3　物理学家的**空间 — 时间**是序偶 (x, t) 的集合, 其中 x 是起点为给定的点 O 的向量, 而 t 是实数, 称为时间. 这是笛卡儿乘积 $\mathrm{M} \times \mathbf{R}$, 其中 M 是通常的空间. 因此看作实向量空间, 空间 — 时间的维数是 $3 + 1 = 4$.

设 L 是一个 n 维向量空间. 在 §16 定理 1 中曾经看到, 它的对偶空间具有由 n 个向量组成的一个基, 因而有

$$\dim(\mathrm{L}) = \dim(\mathrm{L}^*).$$

定理 9　设 L 是一个域上的有限维向量空间, M 是 L 的一个子空间, 而 M^0 是 M 在 L^* 内的零化子, 则有

$$\dim(\mathrm{M}) + \dim(\mathrm{M}^0) = \dim(\mathrm{L}).$$

令 $\dim(\mathrm{M}) = p$, $\dim(\mathrm{L}) = r$. 正如在定理 2 的推论 2 的证明中所看到的, 存在 L 的基 $(x_i)_{1 \leqslant i \leqslant r}$, 使得 $(x_i)_{1 \leqslant i \leqslant p}$ 是 M 的基. 第 2 小节的注 2 表明 L^* 的子空间 M^0 具有由 $r - p$ 个向量组成的基, 这就完成了证明.

推论　设 M 是一个域上的有限维向量空间 L 的一个子空间, 则有

$$\dim(\mathrm{M}) \leqslant \dim(\mathrm{L}),$$

当且仅当 $\mathrm{M} = \mathrm{L}$ 时有 $\dim(\mathrm{M}) = \dim(\mathrm{L})$.

第一个断言直接从定理 9 得到. 这个定理还指出关系 $\dim(\mathrm{M}) = \dim(\mathrm{L})$ 等价于

$$\dim(\mathrm{M}^0) = 0,$$

即 $\mathrm{M}^0 = \{0\}$. 而我们知道 (定理 3 的推论 2) 这个关系等价于 $\mathrm{L} = \mathrm{M}$.

6. 基和维数的特征

下面一个定理提供了有限维向量空间的基的多个有用的特征:

定理 10 设 $x_1, \cdots, x_n$ 是域 K 上有限维向量空间 M 的元素, 则以下性质是等价的:

a) 向量 $x_i (1 \leqslant i \leqslant n)$ 组成 M 的一个基.

b) 向量 x_i 是线性无关的, 并且 M 的维数是 n.

c) 向量 x_i 是线性无关的, 并且 M 的所有自由的子集至多含有 n 个元素.

d) 向量 x_i 生成 M, 并且 M 的维数是 n.

e) 向量 x_i 生成 M, 并且 M 的所有生成元组至少含有 n 个元素.

显然 a) 蕴含 b).

为了证明 b) 蕴含 c), 我们注意一个自由组是 M 的一个基的子集 (定理 2 的推论 1), 从而 (定理 6) 如果 M 的维数是 n, 该子集至多含有 n 个元素.

为了证明 c) 蕴含 d), 考虑一个 $x \in \mathrm{M}$; 根据 c), $n+1$ 个向量 $x_1, \cdots, x_n, x$ 由一个非平凡的线性关系

$$\lambda_1 x_1 + \cdots + \lambda_n x_n + \lambda x = 0$$

相联系. 这里我们有 $\lambda \neq 0$, 否则将会有一个 x_i 的非平凡线性关系, 这与 c) 矛盾, 故 λ 是可逆的, 并且有

$$x = -\lambda^{-1}\lambda_1 x_1 - \cdots - \lambda^{-1}\lambda_n x_n,$$

这就证明了 x_i 生成 M, 从而组成 M 的一个基, 因此 M 的维数是 n.

性质 d) 蕴含 e), 因为 M 的所有生成元组包含 M 的一个基 (定理 2), 从而如果 $\dim(\mathrm{M}) = n$, 则 M 的所有生成元组至少含有 n 个元素.

为了完成证明, 剩下要证的是 e) 蕴含 a). 根据定理 2 可以从族 $(x_i)_{1 \leqslant i \leqslant n}$ 抽取 M 的一个基, 这个基是一个生成元组, 根据 e) 至少含有 n 个元素, 这必定是整个族 $(x_i)_{1 \leqslant i \leqslant n}$. 于是这个族是 M 的一个基, 定理证毕.

定理 11 设 M 是域 K 上的一个向量空间, 而 n 是一个整数, 则下列性质是等价的:

a) M 的维数是 n.

b) n 是从 M 抽取的线性无关的向量的一个族的元素数目中的最大整数.

c) n 是从 M 抽取的生成 M 的向量的一个族的元素数目中的最小整数.

性质 a) 和 b) 的等价性来自定理 10 的性质 a) 和 c) 之间的等价性. 同样, 性质 a) 和 c) 之间的等价性来自定理 10 的性质 a) 和 e) 之间的等价性.

定理 12　域 K 上的向量空间 M 是有限维的, 必须并且只需存在一个整数 n, 使得 M 的所有自由的子集最多具有 n 个元素. 此时有 $\dim(\mathrm{M}) \leqslant n$.

根据前面的结果, 条件显然是必要的. 反之, 如果条件满足, 我们考虑最大的整数 r, 使得 M 包含 r 个线性无关的向量的一个族. 根据假设 $r \leqslant n$, 由定理 11 的 c), M 的维数是 r, 故 M 是有限维的, 并且 $\dim(\mathrm{M}) \leqslant n$.

7. 同态的核与像的维数

当基础环是域时, §18 的定理 1 有如下形式:

定理 13　设 L 和 M 是一个域上的向量空间, 而 f 是从 L 到 M 内的同态. 如果 L 是有限维的, 则有

$$\dim(\mathrm{L}) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)).$$

事实上, $\mathrm{Ker}(f)$ 是有限维的, 因此同构于 K^p, 这里 $p = \dim(\mathrm{Ker}(f))$, 而 $\mathrm{Im}(f)$ 同样是有限维的 (一个有限生成模在一个同态下的像显然是有限生成的), 因此同构于 K^q, 这里 $q = \dim(\mathrm{Im}(f))$. 为了得到定理 13, 只需利用 §18 的定理 1.



注 9　定理 13 虽然是 §18 所证明的结果的一个简单的特殊情形, 但不能够有比这个证明更简单的证明. 反之给以比 §18 的定理 1 更复杂的直接证明是容易的.

我们要讲述的定理 13 的推论在实际中至少与定理 13 本身同样重要, 尤其是

推论 1　设 L 和 M 是一个域上的有限维向量空间, 而 f 是从 L 到 M 内的同态. 假定 $\dim(\mathrm{L}) = \dim(\mathrm{M})$, 那么下列性质是等价的:

a) f 是双射的.

b) f 是满射的.

c) f 是单射的.

d) $\mathrm{Ker}(f) = \{0\}$.

早就已经知道 (§7, 定理 8) c) 和 d) 是等价的. 由于 a) 是与 b) 和 c) 同时成立的, 一切都归结为证明 b) 和 d) 的等价性. 而 b) 等价于关系 (定理 9 的推论)

$$\dim(\mathrm{Im}(f)) = \dim(\mathrm{M}), \tag{12}$$

而 d) 等价于

$$\dim(\mathrm{Ker}(f)) = 0; \tag{13}$$

于是 b) 和 d) 的等价性由定理 13 和条件 $\dim(\mathrm{L}) = \dim(\mathrm{M})$ 得到, 故得推论.

当 $\mathrm{L} = \mathrm{M}$ 时最经常使用的是推论 1. 下一节将看到还可以把它翻译成线性方程组的语言.

推论 2 设 E 和 F 是一个域上的有限维向量空间 M 的向量子空间, 则有

$$\dim(\mathrm{E}+\mathrm{F}) = \dim(\mathrm{E}) + \dim(\mathrm{F}) - \dim(\mathrm{E}\cap\mathrm{F}).$$

事实上, 考虑由 $f(x,y) = x+y$ 定义的同态 $f : \mathrm{E}\times\mathrm{F} \to \mathrm{M}$ (§17, 第 3 小节). 我们有 E+F=Im(f), 此外 Ker(f) 同构于 $\mathrm{E}\cap\mathrm{F}$ (§17, 注 1), 故得

$$\begin{aligned}\dim(\mathrm{E}) + \dim(\mathrm{F}) &= \dim(\mathrm{E}\times\mathrm{F}) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) \\ &= \dim(\mathrm{E}\cap\mathrm{F}) + \dim(\mathrm{E}+\mathrm{F}).\end{aligned}$$

这就是所要证明的关系.

推论 3 设 $\mathrm{E}_1, \cdots, \mathrm{E}_r$ 是有限维向量空间 E 的向量子空间, 则有

$$\dim(\mathrm{E}_1+\cdots+\mathrm{E}_r) \leqslant \dim(\mathrm{E}_1) + \cdots + \dim(\mathrm{E}_r),$$

为了这个关系的两端相等, 必须并且只需子空间 $\mathrm{E}_1, \cdots, \mathrm{E}_r$ 是线性无关的.

事实上, 考虑由

$$f(x_1, \cdots, x_r) = x_1 + \cdots + x_r$$

定义的同态

$$f : \mathrm{E}_1\times\cdots\times\mathrm{E}_r \to \mathrm{E}_1+\cdots+\mathrm{E}_r.$$

它是满射的, 因此从定理 13 得到

$$\begin{aligned}\dim(\mathrm{E}_1+\cdots+\mathrm{E}_r) &= \dim(\mathrm{E}_1\times\cdots\times\mathrm{E}_r) - \dim(\mathrm{Ker}(f)) \\ &= \dim(\mathrm{E}_1) + \cdots + \dim(\mathrm{E}_r) - \dim(\mathrm{Ker}(f)).\end{aligned}$$

由此得到所要证明的不等式. 而要等式成立, 当且仅当 $\mathrm{Ker}(f) = 0$, 而这正表明 (§17, 第 3 小节) E 的子空间 E_i 是线性无关的.

8. 同态、向量族和矩阵的秩

设 L 和 M 是有限维向量空间, 而 f 是从 L 到 M 内的一个同态. 称 M 的子空间 $\mathrm{Im}(f) = f(\mathrm{L})$ 的维数为**同态 f 的秩**.

设 $(a_i)_{1\leqslant i\leqslant n}$ 是 L 的一个生成元组, 那么 Im(f) 是由向量 $x_i = f(a_i)$ 生成的. 这就引导我们到以下定义: 给定向量空间 M (有限维的或无限维的) 的元素的一个族 $(x_i)_{1\leqslant i\leqslant n}$, 称由 x_i 生成的 M 的子空间的维数 (注意到即使 M 是无限维的, 这个子空间也是有限维的) 为**族 (x_i) 的秩**.

设 r 是向量空间 M 的元素的一个族 $(x_i)_{1\leqslant i\leqslant n}$ 的秩, 并且用 M′ 表示由 x_i 生成的 M 的子空间. 根据定理 2, 从族 $(x_i)_{1\leqslant i\leqslant n}$ 可以抽取出 M′ 的一个基, 这个基必定含有 r 个元素. 如果有需要, 修改 x_i 的编号, 可以假定 $x_1,\cdots,x_r$ 是 M′ 的一个基. 因此有 $x_1,\cdots,x_r$ 是线性无关的, 并且 M′ 的所有元素是这些元素的线性组合, 即有关系

$$x_j=\rho_{j1}x_1+\cdots+\rho_{jr}x_r \quad 对于\ r+1\leqslant j\leqslant n. \tag{14}$$

反之, 设有关系 (14), 由 $x_1,\cdots,x_r$ 生成的子空间不仅含有这些向量, 而且含有 $x_{r+1},\cdots,x_n$, 从而 $x_1,\cdots,x_r$ 生成 M′. 如果 $x_1,\cdots,x_r$ 是线性无关的, 它们就组成 M′ 的一个基, 从而 M′ 的维数是 r, 因此族 $(x_i)_{1\leqslant i\leqslant n}$ 的秩是 r.

这些考虑自然应用到 M 是一个向量空间 L 的对偶空间, 这时 x_i 就是 L 上的线性型 f_i, 这就允许定义 L 上的**一族线性型的秩**. 显然这个概念隐含在定理 5 的陈述中.

现在设

$$\mathrm{A}=(\alpha_{ij})_{1\leqslant i\leqslant n,1\leqslant j\leqslant p}$$

是元素在域 K 内的一个矩阵. 称由 A 表示的从 K^p 到 K^n 的同态 (§12, 第 3 小节, 注 1) 的秩为**矩阵 A 的秩**. 这个同态的像是由 K^p 的典范基的元素的像, 即矩阵 A 的列向量, 生成的 K^n 的向量子空间. 故矩阵 A 的秩是由矩阵 A 的 p 个列生成的 K^n 的向量子空间的维数.

定理 14 设 L 和 M 是有限维向量空间, 而 f 是从 L 到 M 内的一个同态, 而 A 是 f 关于 L 的一个基 $(a_i)_{1\leqslant i\leqslant p}$ 和 M 的一个基 $(b_j)_{1\leqslant j\leqslant n}$ 的矩阵, 则 f 的秩等于 A 的秩.

事实上, 考虑同构

$$u:\mathrm{K}^p\to\mathrm{L},\quad v:\mathrm{K}^n\to\mathrm{M},$$

它们分别映射 K^p 和 K^n 的典范基到 L 和 M 的给定的基, 那么同态

$$v^{-1}\circ f\circ u:\mathrm{K}^p\to\mathrm{K}^n$$

正好以 A 作为关于 K^p 和 K^n 的典范基的矩阵 (§12, 第 3 小节, 注 2), 根据定义这个同态的秩就是矩阵 A 的秩. 但由于 u 和 v 是同构的, 显然 f 的像和 $v^{-1}\circ f\circ u$ 的像是同构的, 由此得到这两个同态有同样的秩. 故得定理.

定理 15 设 M 是有限维向量空间, $(b_j)_{1\leqslant j\leqslant n}$ 是 M 的一个基, 并且

$$x_i=\alpha_{1i}b_1+\cdots+\alpha_{ni}b_n \quad (1\leqslant i\leqslant p)$$

是 M 的元素的一个族, 那么族 $(x_i)_{1\leqslant i\leqslant p}$ 的秩等于矩阵 $(\alpha_{ij})_{1\leqslant i\leqslant n,1\leqslant j\leqslant p}$ 的秩.

为了证明定理只需应用前一个定理, 其中取 L 为空间 K^p, 取 K^p 的典范基为 L 的基, 把典范基映射到给定向量 x_i 的同态取作 f, x_i 生成子空间 $\mathrm{Im}(f)$, 以致族 (x_i) 的秩等于 f 的秩, 即 f 的矩阵的秩, 而该矩阵正是矩阵 (α_{ij}).

9. 矩阵的秩的计算

前一个小节的结果表明同态或向量族的秩的计算可以归结为矩阵的秩的计算. 这一主题的基本定理是下列定理:

定理 16 设 $\mathrm{A} = (\alpha_{ij})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant p}$ 是元素在域 K 内的一个矩阵, 则 A 的秩是能从 A 抽出 r 阶可逆方阵的 r 中的最大整数.

设 L 和 M 是 K 上的维数分别为 p 和 n 的向量空间, 选择 L 的基 $(a_i)_{1\leqslant i\leqslant p}$ 和 M 的基 $(b_j)_{1\leqslant j\leqslant n}$, 并且考虑从 L 到 M 内的同态 f, 使得 f 关于所选择的基的矩阵正好是 A. A 的秩等于 f 的秩, 即 $\mathrm{Im}(f)$ 的维数. 设这个秩是 r, 我们首先证明实际上可以从 A 抽取一个可逆的 r 阶方阵.

根据定理 13, 我们有 $\dim(\mathrm{Ker}(f)) = p - r$, 因此 $\mathrm{Ker}(f)$ 具有 $p-r$ 个向量 $c_1, \cdots, c_{p-r}$ 组成的一个基. 由于向量

$$c_1, \cdots, c_{p-r}, a_1, \cdots, a_p$$

生成 L, 并且它们中的前 $p-r$ 个向量线性无关, 就可以将 $p-r$ 个向量 $c_1, \cdots, c_{p-r}$ 添加在 $a_1, \cdots, a_p$ 中选择的 r 个向量, 使它们组成 L 的基. 如果有必要, 修改向量 $a_1, \cdots, a_p$ 的编号 (这意味着置换矩阵 A 的列), 可以假定

$$c_1, \cdots, c_{p-r}, a_1, \cdots, a_r$$

组成 L 的一个基. 用 L' 表示由 $a_1, \cdots, a_r$ 生成的 L 的子空间, 则有

$$\mathrm{L} = \mathrm{Ker}(f) \oplus \mathrm{L}'.$$

设 $x = y + z$, 其中 $y \in \mathrm{Ker}(f)$, $z \in \mathrm{L}'$, 显然有 $f(x) = f(z)$, 于是 $f(\mathrm{L}') = \mathrm{Im}(f)$. 而由于 $\mathrm{Ker}(f) \cap \mathrm{L}' = \{0\}$, f 在 L' 上的限制是从 L' 到 M 的子空间 $\mathrm{Im}(f)$ 上的同构. 由此得到向量

$$f(a_1), \cdots, f(a_r)$$

组成 $\mathrm{Im}(f)$ 的一个基.

像上面那样, 重新利用定理 2 我们发现, 如果有必要调整 b_j 的编号, 可以假定向量

$$f(a_1), \cdots, f(a_r), b_{r+1}, \cdots, b_n$$

组成 M 的一个基. 于是有

$$\mathrm{M} = \mathrm{Im}(f) \oplus \mathrm{M}'', \tag{15}$$

其中 M″ 是由 $b_{r+1},\cdots,b_n$ 生成的 M 的子空间. 此外还有分解

$$\mathrm{M}=\mathrm{M}'\oplus\mathrm{M}'', \tag{16}$$

其中的 M′ 是 $b_1,\cdots,b_r$ 生成的线性空间. 用 u 表示直和分解 (16) 所确定的从 M 到 M′ 上的同态, 即如果

$$x=x'+x'',\quad \text{其中 } x'\in\mathrm{M}',\ x''\in\mathrm{M}'',$$

则令 $u(x)=x'$ (§17, 第 4 小节).

显然 $\mathrm{Ker}(u)=\mathrm{M}''$, 而由于 (15) 是直和, 故有

$$\mathrm{Ker}(u)\cap\mathrm{Im}(f)=\{0\},$$

于是由 u 诱导的从 $\mathrm{Im}(f)$ 到 M′ 内的映射是单射的; 但由于

$$\dim(\mathrm{Im}(f))=\dim(\mathrm{M}')=r,$$

这个映射事实上是双射的 (定理 13 的推论 1). 由于我们已经证明 f 诱导一个从 L′ 到 $\mathrm{Im}(f)$ 上的双射, 故由

$$f'(x)=u(f(x))\quad \text{对于所有 } x\in\mathrm{L}'$$

定义的映射

$$f':\mathrm{L}'\to\mathrm{M}'$$

是一个双射, 即是一个同构, 于是 f' 关于 L′ 的基 $(a_1,\cdots,a_r)$ 和 M′ 的基 $(b_1,\cdots,b_r)$ 的矩阵是可逆的 (§15, 第 1 小节). 而如果 $\mathrm{A}=(\alpha_{ij})_{1\leqslant i\leqslant n,1\leqslant j\leqslant p}$, 则有

$$f(a_i)=\alpha_{1i}b_1+\cdots+\alpha_{ni}b_n,$$

因此有

$$u(f(a_i))=\alpha_{1i}b_1+\cdots+\alpha_{ri}b_r,$$

从而 f' 的矩阵就是由 A 的前 r 行和前 r 列组成的矩阵 $\mathrm{A}=(\alpha_{ij})_{1\leqslant i,j\leqslant r}$, 而这就如所宣布的那样证明了如果 A 的秩是 r, 就可以从 A 抽取一个 r 阶可逆的方阵.



注 10　一般来说, 并不是矩阵 $(\alpha_{ij})_{1\leqslant i,j\leqslant r}$ 是可逆的. 不要忘记事实上在前面的证明过程中, 应当允许矩阵 A 的行的置换和列的置换.

为了完成定理的证明, 还要证明如果 A 的秩是 r, 并且能够从 A 抽取一个 s 阶的可逆方阵, 则必然有 $s\leqslant r$.

比如说假定 $(\alpha_{ij})_{1\leqslant i,j\leqslant s}$ 是可逆的. 现在用 L′ 表示由 $a_1,\cdots,a_s$ 生成的 L 的子空间, 用 M′ 表示由 $b_1,\cdots,b_s$ 生成的 M 的子空间, 用 M″ 表示由 $b_{s+1},\cdots,b_n$ 生成

的 M 的子空间, 用 u 表示从 M 到 M′ 上的平行于 M″ 的投影. 那么 $(\alpha_{ij})_{1\leqslant i,j\leqslant s}$ 是关于基 $(a_i)_{1\leqslant i\leqslant s}$ 和 $(b_j)_{1\leqslant j\leqslant s}$ 的、由 $f'(x)=u(f(x))$ 定义的从 L′ 到 M′ 的同态 f' 的矩阵. 根据假设这个矩阵是可逆的, 故 f' 是双射的, 从而是单射的, f 在 L′ 的限制更加是单射的, 从而映射 L′ 到一个与 L′ 同样维数的子空间上, 这就表明 Im(f) 至少有维数 s. 换句话说, $s\leqslant r$, 这就结束了证明.

注 11　如果域 K 是交换的, 我们已经准备好了一个理论上十分简单的、决定一个方阵是否可逆的准则: 只需检查它的行列式 (§23, 定理 8 的推论 1).

例 4　考虑 K^3 的两个向量

$$x'=(a',b',c'),\quad x''=(a'',b'',c''),$$

说它们是线性无关的, 就是说矩阵

$$\begin{pmatrix} a' & a'' \\ b' & b'' \\ c' & c'' \end{pmatrix}$$

的秩是 2, 也就是矩阵

$$\begin{pmatrix} b' & b'' \\ c' & c'' \end{pmatrix},\quad \begin{pmatrix} a' & a'' \\ c' & c'' \end{pmatrix},\quad \begin{pmatrix} a' & a'' \\ b' & b'' \end{pmatrix}$$

中至少一个是可逆的, 这就是说

$$b'c''-b''c',\quad a'c''-a''c',\quad a'b''-a''b'$$

不全是零 (§15, 第 3 小节).

推论　*元素在一个域内的矩阵 A 的秩等于转置矩阵 ${}^tA$ 的秩.*

事实上, 显然从 ${}^tA$ 抽出的方阵是从 A 抽出的方阵的转置. 如果一个矩阵是可逆的, 则其转置矩阵也是可逆的, 反之亦然.

10. 从其方程计算向量子空间的维数

设 L 是域 K 上的 n 维向量空间, 而 M 是 L 的一个向量子空间. 前面曾经讲过 (定理 3 的推论 1) 存在 L 上的线性型 $f_i(1\leqslant i\leqslant m)$, 使得 M 是满足关系

$$f_1(x)=\cdots=f_m(x)=0 \tag{17}$$

的 $x\in L$ 的集合. 反之, 显然对于任何线性型 f_i, 关系 (17) 定义 L 的一个向量子空间 M.

这一点交代清楚后, M 的维数的计算可以这样实现:

定理 17 设 $f_1, \cdots, f_m$ 是域 K 上 n 维向量空间 L 上的线性型. 由关系 (17) 定义的 L 的向量子空间 M 的维数等于 $n-r$, 这里 r 是族 $(f_i)_{1 \leqslant i \leqslant m}$ 的秩.

可以假定 $f_1, \cdots, f_r$ 是线性无关的, 在这一条件下, 有关系

$$f_j = \rho_{j1} f_1 + \cdots + \rho_{jr} f_r \quad \text{对于 } r+1 \leqslant j \leqslant m,$$

那么显然 (17) 的解和

$$f_1(x) = \cdots = f_r(x) = 0 \tag{18}$$

的解相同. 这样就把问题归结为给定的线性型 f_i 线性无关的情形.

考虑由

$$f(x) = (f_1(x), \cdots, f_r(x))$$

给定的同态 $f: \mathrm{L} \to \mathrm{K}^r$. 显然 $\mathrm{M} = \mathrm{Ker}(f)$, 并且因此 (定理 13) 有

$$\dim(\mathrm{M}) = n - \dim(\mathrm{Im}(f)).$$

最后问题归结为证明

$$\dim(\mathrm{Im}(f)) = r = \dim(\mathrm{K}^r),$$

即 $\mathrm{Im}(f) = \mathrm{K}^r$, 而这可以从定理 4 推出.

§19 习题

1. 考虑 $\mathbf{R}^4$ 中的向量

$$(1,0,0,-1), \quad (2,1,1,0), \quad (1,1,1,1), \quad (1,2,3,4), \quad (0,1,2,3),$$

从这五个向量中抽取它们生成的子空间的一个基.

2. 在 $\mathbf{R}^4$ 内考虑由向量

$$(1,1,1,1), \quad (1,-1,1-1), \quad (1,3,1,3)$$

生成的子空间 L 和由

$$(1,2,0,2), \quad (1,2,1,2), \quad (3,1,3,1)$$

生成的子空间 M. 计算 $\mathrm{L} \cap \mathrm{M}$ 和 $\mathrm{L} + \mathrm{M}$ 的维数.

3. 设 V 是一个域上的 n 维向量空间, 而 L 是 V 的一个 r 维子空间. V 的一个子空间 M 是 L 在 V 内的补空间, 必须并且只需有

$$\mathrm{L} \cap \mathrm{M} = \{0\}, \quad \dim(\mathrm{M}) = n - r.$$

4. 设 V 是一个域上的有限维向量空间, $(a_i)_{1 \leqslant i \leqslant n}$ 是 V 的一个基, $x_1, \cdots, x_r$ 是 V 的元素, $r \leqslant n$. 诸 x_i 是线性无关的, 必须并且只需存在 V 的一个同构 u, 使得

$$u(a_i) = x_i \quad \text{对于 } 1 \leqslant i \leqslant r.$$

¶5. 设 L 和 M 是域 K 上的有限维向量空间, 而 f 是从 L 到 M 内的秩为 r 的一个同态. 证明存在 L 的一个基 $(a_i)_{1\leqslant i\leqslant p}$ 和 M 的一个基 $(b_j)_{1\leqslant j\leqslant q}$, 使得有关系

$$f(a_i) = b_i \quad 对于\ 1 \leqslant i \leqslant r,$$
$$f(a_i) = 0 \quad 对于\ r+1 \leqslant i \leqslant p.$$

由此推出下列结果: 设 A 是一个其元素在 K 内的 q 行 p 列的秩为 r 的矩阵, 则存在矩阵

$$\mathrm{U} \in \mathrm{GL}(q, \mathrm{K}) \quad 和 \quad \mathrm{V} \in \mathrm{GL}(p, \mathrm{K}),$$

使得

$$\mathrm{UAV} = \begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$$

(出现在右端的 0 表示零矩阵, 它们的行数和列数使得右端是 q 行 p 列矩阵).

设 A 和 B 是元素在 K 内的 q 行 p 列的两个矩阵. 存在矩阵 $\mathrm{U} \in \mathrm{GL}(q, \mathrm{K})$ 和 $\mathrm{V} \in \mathrm{GL}(p, \mathrm{K})$, 使得 $\mathrm{B} = \mathrm{UAV}$, 必须并且只需 A 和 B 有同样的秩.

6. 设 E 和 F 是有限维向量空间 V 的两个子空间. 存在 V 的一个使得 $\mathrm{F} = u(\mathrm{E})$ 的自同构 u, 必须并且只需 $\dim(\mathrm{E}) = \dim(\mathrm{F})$.

7. 设 $x_1, \cdots, x_n$ 是一个向量空间的元素. 假定 $x_1, \cdots, x_r$ 是线性无关的, 并且有关系

$$x_j = \rho_{j1}x_1 + \cdots + \rho_{jr}x_r \quad (r+1 \leqslant j \leqslant n).$$

设 L 是 $x_1, \cdots, x_n$ 之间的线性关系的空间, 即使得

$$\lambda_1 x_1 + \cdots + \lambda_n x_n = 0$$

的标量组 $(\lambda_1, \cdots, \lambda_n) \in \mathrm{K}^n$ 的集合. 证明 $n-r$ 个元素

$$(\rho_{j1}, \cdots, \rho_{jr}, 0, \cdots, 0, -1, 0, \cdots, 0)$$

(其中的 -1 在第 j 个位置) 组成 L 的一个基.

8. 设 $f_1, \cdots, f_n$ 是一个向量空间上的线性型. 方程组

$$f_i(x) = \beta_i \quad (1 \leqslant i \leqslant n)$$

至少具有一个解, 必须并且只需对于 $f_1, \cdots, f_n$ 之间的所有线性关系 $(\lambda_1, \cdots, \lambda_n)$ 有

$$\lambda_1\beta_1 + \cdots + \lambda_n\beta_n = 0.$$

9. 设 K 是一个交换域.

a) 把 $\mathrm{M}_n(\mathrm{K})$ 看作 K 上的 n^2 维向量空间, $\mathrm{M}_n(\mathrm{K})$ 上的一个线性型 f 对于任意 $\mathrm{X}, \mathrm{Y} \in \mathrm{M}_n(\mathrm{K})$ 满足

$$f(\mathrm{XY}) = f(\mathrm{YX}),$$

必须并且只需存在一个标量 $\lambda \in \mathrm{K}$, 使得

$$f(\mathrm{X}) = \lambda \cdot \mathrm{Tr}(\mathrm{X}) \quad 对于所有\ \mathrm{X} \in \mathrm{M}_n(\mathrm{K})$$

(我们提醒 (见 §12, 习题 8) Tr(X) 表示 X 的对角线元素之和).

b) 由此和 §19 的定理 3 推出以下结果: 一个矩阵 $A \in M_n(K)$ 可以写成形式 $XY - YX$, 必须并且只需 $Tr(A) = 0$.

10. 设 V 是域 K 上的有限维向量空间, 而 $(a_i)_{1\leqslant i\leqslant n}$ 是 V 的一个基.

a) 设 x 是 V 的一个非零元. 证明 (利用 §19 的结果) 存在一个指标 i, 使得 $a_1, \cdots, a_{i-1}, x, a_{i+1}, \cdots, a_n$ 还组成 V 的一个基.

b) 设 $x_1, \cdots, x_p$ 是 V 内的线性无关的向量. 借助问题 a) 并且关于 p 进行归纳推理, 证明存在 V 的由 p 个向量 x_i 和 $n-p$ 个向量 a_i 组成的一个基. 由此推出 §19 的定理 6 和 7 的一个新证明.

11. 关于 p 进行归纳推理直接证明定理 7 的推论 (利用方程组中一个方程把一个未知元用其余的 $p-1$ 个未知元表示, 并且归结到 $p-1$ 个未知元的 $n-1$ 个方程的方程组).

¶12. 设 L 和 M 是一个域上的两个有限维向量空间, f 是从 L 到 M 内的一个线性映射, 并且

$$^t f : M^* \to L^*$$

是其转置映射 (§16). 证明 $Im(^t f)$ 是 $Ker(f)$ 的零化子, 并且 $Ker(^t f)$ 是 $Im(f)$ 的零化子 (利用定理 3). 由此推出 f 和它的转置有同样的秩, 并且应用这个结果得到定理 16 的推论的另一个证明. 证明如果 f 是 L 的一个自同态, 则 f 和 $^t f$ 的核有同样的维数.

13. 设 K 是一个域, 一个矩阵 $A \in M_n(K)$ 是可逆的, 必须并且只需 A 不是环 $M_n(K)$ 内的零因子.

14. 设 X 是一个其元素在一个域内的矩阵, 证明如果 X 的一行 (对应的, 列) 加上其余的行 (对应的, 列) 的一个线性组合, 或者对于行 (对应的, 列) 进行任意置换, 则 X 的秩不变.

15. 设 V 是一个有限维复向量空间, 而 $(a_k)_{1\leqslant k\leqslant n}$ 是 V 的一个基. 把 V 看作一个实向量空间, 证明 $2n$ 个向量

$$a_1, \cdots, a_n, \quad ia_1, \cdots, ia_n$$

组成 **R** 上的 V 的一个基. 由此推出

$$\dim_{\mathbf{R}}(V) = 2 \cdot \dim_{\mathbf{C}}(V).$$

¶16. 设 L 是一个域, 而 K 是 L 的一个子域, 称 L 是 K 的一个**有限次扩张**, 如果看作 K 上的向量空间的 L 是有限维的, 其维数称为 L 在 K 上的**次数**, 记为 [L : K].

设 V 是 L 上的有限维向量空间, 把 V 看作 K 上的向量空间. 证明, 如果 L 是 K 的一个有限扩张, 则有

$$\dim_K(V) = [L : K] \cdot \dim_L(V)$$

(模仿前一个习题的推理).

假定 K 是一个域, L 是 K 的一个有限扩张域, 而 M 是 L 的一个有限扩张域. 证明 M 是 K 的一个有限次扩张, 并且

$$[M : K] = [M : L] \cdot [L : K].$$

17. 设 $L_1 \xrightarrow{f_1} L_2 \xrightarrow{f_2} L_3 \cdots \xrightarrow{f_n} L_{n+1}$ 是由一个域上的有限维向量空间和从一个空间到下一个内的同态组成的序列. 假定 f_1 是单射, f_n 是满射, 并且

$$Im(f_i) = Ker(f_{i+1}) \quad 对于 \ 1 \leqslant i \leqslant n-1.$$

证明有

$$\dim(L_1) - \dim(L_2) + \dim(L_3) - \cdots + (-1)^n \dim(L_{n+1}) = 0.$$

¶18. 设 p 是一个素数. 为了能够解同余方程组

$$\begin{cases} x + 2y + 3z \equiv u \pmod p, \\ 4x + 5y + 6z \equiv v \pmod p, \end{cases}$$

整数 u 和 v 应当满足什么条件?

¶19. 设 V 是一个域上的 n 维向量空间.

a) 设 $L_1, \cdots, L_r$ 是 V 的向量子空间, 满足条件

$$L_1 \subset L_2 \subset \cdots \subset L_r = V,$$

令 $\dim(L_i) = d_i$. 证明存在 V 的一个基, 使得对于每个 i, 这个基的前 d_i 个元素组成 L_i 的一个基.

b) 设 u 是 V 的一个幂零同态, 即对于至少一个正整数 k 有 $u^k = 0$. 设 r 是使得

$$u^r = 0$$

的最小正整数. 令

$$L_1 = \mathrm{Ker}(u), \quad L_2 = \mathrm{Ker}(u^2), \cdots, L_r = \mathrm{Ker}(u^r).$$

证明 V 的这些子空间满足问题 a) 的假设. 由此推出存在 V 的一个基, 使得 u 关于这个基的矩阵有形式(*)

$$\begin{pmatrix} 0 & * & * & * & \cdots & * \\ 0 & 0 & * & * & \cdots & * \\ 0 & 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

其逆成立吗?

c) 设 $N \in M_n(K)$ 是一个幂零矩阵. 证明存在一个矩阵 $U \in GL(n, K)$, 使得

$$UNU^{-1} = \begin{pmatrix} 0 & * & * & * & \cdots & * \\ 0 & 0 & * & * & \cdots & * \\ 0 & 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

其逆成立吗?

¶¶20. 设 V 是 $\mathbf{Q}$ 上的一个 n 维向量空间, 而 M 是 V 的有限生成子群.

a) 证明存在 V 的一个基, 所有 $x \in M$ 关于这个基的坐标是整数.

(*) 在 §35 将找到更精确的结果.

b) 证明存在一个整数 $r \leqslant n$, 使得 M 是由 V 的 r 个适当选择的线性无关向量生成的 V 的子空间 (换句话说, 对于一个整数 $r \leqslant n$, 群 M 同构于 $\mathbf{Z}^r$). (利用 §18 的定理 3.)

c) 设 $(a_i)_{1\leqslant i\leqslant n}$ 是 V 的一个基. 证明存在 V 的一个基 $(b_i)_{1\leqslant i\leqslant n}$, 使得 (i) M 是由 $b_1, \cdots, b_r$ 生成的; (ii) 从基 $(a_i)_{1\leqslant i\leqslant n}$ 到 $(b_i)_{1\leqslant i\leqslant n}$ 的过渡矩阵是三角矩阵 (模仿 §18 定理 3 的证明).

d) 由此推出所有矩阵 $\mathrm{X} \in \mathrm{M}_n(\mathbf{Q})$ 是一个矩阵 $\mathrm{U} \in \mathrm{GL}(n, \mathbf{Z})$ 和一个三角矩阵的乘积.

¶21. (Sylvester 不等式) 设 A 和 B 是其元素在一个域内的两个 n 阶方阵. 证明

$$\mathrm{rank}(\mathrm{A}) + \mathrm{rank}(\mathrm{B}) - n \leqslant \mathrm{rank}(\mathrm{AB}) \leqslant \min(\mathrm{rank}(\mathrm{A}), \mathrm{rank}(\mathrm{B})).$$

22. 设 u 是一个交换域 K 上的有限维向量空间 V 的一个自同态. 证明, 如果 A 和 B 是 u 关于 V 的任意两个基的矩阵, 则有

$$\mathrm{Tr}(\mathrm{A}) = \mathrm{Tr}(\mathrm{B})$$

(§12, 习题 8). u 关于 V 的不同的基的矩阵的迹的公共值称为**自同态 u 的迹**, 记作

$$\mathrm{Tr}(u).$$

证明

$$\mathrm{Tr}(u+v) = \mathrm{Tr}(u) + \mathrm{Tr}(v), \quad \mathrm{Tr}(\lambda u) = \lambda \mathrm{Tr}(u) \quad \text{对于 } \lambda \in \mathrm{K},$$
$$\mathrm{Tr}(u \circ v) = \mathrm{Tr}(v \circ u), \quad \mathrm{Tr}(j_v) = \dim(\mathrm{V}),$$

并且这些性质完全刻画了映射

$$\mathrm{Tr} : \mathscr{L}(\mathrm{V}) \to \mathrm{K}$$

的特征. 证明对于 V 的所有自同态还有

$$\mathrm{Tr}(u) = \mathrm{Tr}({}^t u).$$

¶23. 对于系数在一个域内的线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_1 \end{cases}$$

至少有一个解, 必须并且只需两个矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad \text{和} \quad \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}$$

具有同样的秩.

¶¶24. 设 L 是一个交换域, 而 A 是 L 的一个子环, 假定作为 A- 模, L 是有限生成的. 现在要由此推出 A 是 L 的一个子域.

a) 设 K 是这样的 $x \in \mathrm{L}$ 的集合, 它可以写成形式 u/v, 其中的 $u, v \in \mathrm{A}$, 并且 $v \neq 0$. 证明这是 L 的一个子域, 并且 L 作为 K 上的向量空间是有限维的.

b) 设 $(u_1, \cdots, u_n)$ 是 A-模 L 的一组生成元. 选择作为 K 上的向量空间 L 的一个基 $(v_0, v_1, \cdots, v_r)$, 其中的 $v_0 = 1$, 令

$$u_i = \sum_{0 \leqslant j \leqslant r} x_{ij} v_j,$$

其中的 $x_{ij} \in \mathrm{K}$. 证明 A-模 K 是由元素 x_{i0} 生成的.

c) 借助 §11 的习题 9 完成证明.

§20 线性方程组

1. 记号和术语

给定一个环 K, 称所有关系组

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{1p}\xi_p = \beta_1, \\ \qquad\cdots\cdots\cdots\cdots \\ \alpha_{n1}\xi_1 + \cdots + \alpha_{np}\xi_p = \beta_n \end{cases} \tag{1}$$

为**系数在 K 内的 p 个未知元 n 个线性方程的方程组**, 其中的**系数** α_{ij} 和*右端*的常数项 β_j 是 K 内的给定元素. 称其坐标满足关系 (1) 的所有向量

$$x = (\xi_1, \cdots, \xi_p) \in \mathrm{K}^p$$

为 (1) 的**解**, 或当要排除可能的混淆时称为在**环 K 内的解**.

为了研究方程组 (1), 在以下提供的三种观点下予以考虑是有用的, 甚至是必须的.

首先, 应当引进关于典范基的矩阵

$$\mathrm{A} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1p} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{np} \end{pmatrix} \tag{2}$$

的同态

$$f : \mathrm{K}^p \to \mathrm{K}^n,$$

并且考虑向量

$$b = (\beta_1, \cdots, \beta_n) \in \mathrm{K}^n,$$

那么根据 §12 第 3 小节, (1) 的解显然正是满足关系

$$f(x) = b \tag{3}$$

的向量 $x \in \mathrm{K}^p$.

其次, 把向量 x 和 b 等同于列矩阵

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_p \end{pmatrix}, \quad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix},$$

那么方程组 (1) 取形式

$$\mathrm{A}x = b. \tag{4}$$

最后, 引进 K^p 上的线性型

$$f_i(\xi_1, \cdots, \xi_p) = \alpha_{i1}\xi_1 + \cdots + \alpha_{ip}\xi_p,$$

那么方程组 (1) 书写为

$$\begin{cases} f_1(x) = \beta_1, \\ \cdots\cdots\cdots\cdots \\ f_n(x) = \beta_n. \end{cases} \tag{5}$$

在本节后面我们使用上面引进的记号而不再加以说明, 并且假定 K 是一个域.

2. 线性方程组的秩, 解的存在性条件

称线性型的族 $(f_j)_{1\leqslant j\leqslant n}$ 的秩 r 为**方程组** (1) **的秩**. 由于 f_j 关于 K^p 的典范基在 $(\mathrm{K}^p)^*$ 内的对偶基的坐标是标量 α_{ij}, 我们发现 (§19, 定理 15) 要找的秩 r 正是 A 的转置矩阵

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{n1} \\ \vdots & & \vdots \\ \alpha_{1p} & \cdots & \alpha_{np} \end{pmatrix}$$

的秩, 故 (§19, 定理 16 的推论) 方程组 (1) 的秩等于矩阵 A 的秩.

设这个秩是 r. 如果有必要调整方程组 (1) 中的编号, 可以假定线性型 $f_1, \cdots, f_r$ 是线性无关的, 于是有关系

$$f_j = \rho_{1j}f_1 + \cdots + \rho_{rj}f_r \quad (r+1 \leqslant j \leqslant n),$$

像在 §19 第 3 小节证明过的那样, 方程组 (5) 具有解, 当且仅当相容性条件

$$\beta_j = \rho_{1j}\beta_1 + \cdots + \rho_{rj}\beta_r \quad (r+1 \leqslant j \leqslant n)$$

满足. 如果这个条件满足, (5) 的解和

$$\begin{cases} f_1(x) = \beta_1, \\ \cdots\cdots\cdots\cdots \\ f_r(x) = \beta_r \end{cases} \tag{6}$$

的解相同. 这就归结为研究左端是线性无关的线性型的方程组. 在 §19 (定理 4) 已经指出这样的方程组至少有一个解.

3. 相伴齐次方程组

如果方程 (3) 至少有一个解 x, 那么其他的解显然有形式 $x+y$, 其中的 y 满足 $f(y)=0$, 即 $y\in\mathrm{Ker}(f)$. 于是 (3) 的解的个数的问题依赖于方程 $f(y)=0$ (参见 §7, 定理 8).

如果令 $y=(\eta_1,\cdots,\eta_p)$, 那么显然方程 $f(y)=0$ 写成

$$\begin{cases}\alpha_{11}\eta_1+\cdots+\alpha_{1p}\eta_p=0,\\ \qquad\cdots\cdots\cdots\cdots\\ \alpha_{n1}\eta_1+\cdots+\alpha_{np}\eta_p=0.\end{cases}\tag{1'}$$

称 (1′) 是给定**方程组** (1) **的相伴线性齐次方程组**; 称解

$$\eta_1=\cdots=\eta_p=0$$

为这个方程组的**平凡解**.

(1′) 还可以写成形式

$$\begin{cases}f_1(x)=0,\\ \cdots\cdots\cdots\cdots\\ f_n(x)=0.\end{cases}\tag{5'}$$

这个方程组的解组成 K^p 的一个向量子空间, 根据 §19 的定理 17, 其维数是 $p-r$, r 是族 $(f_i)_{1\leqslant i\leqslant n}$ 的秩.

4. Cramer 方程组

再提醒一下, 直到本节末, 假设 K 是一个域.

定理 1 以下性质等价:

a) 对于任意标量 $\beta_1,\cdots,\beta_n\in\mathrm{K}$, 方程组 (1) 有且仅有一个解.

b) $p=n$, 并且与方程组 (1) 相伴的齐次方程组只有平凡解.

c) $p=n$, 并且线性型 $f_1,\cdots,f_n$ 是线性无关的.

性质 a) 说明同态

$$f:\mathrm{K}^p\to\mathrm{K}^n$$

是双射. 如果这样, 那么 f 是一个同构, 因此 $\mathrm{Ker}(f)=\{0\}$, 并且

$$\dim(\mathrm{K}^p)=\dim(\mathrm{K}^n),$$

即 $p=n$; 于是 a) 蕴含 b). 反过来, 如果 f 是单射, 并且 $p=n$, 那么根据 §19, 定理 13 的推论 1, f 是双射, 故 b) 蕴含 a).

现在证明 a) 蕴含 c). 已经知道 a) 蕴含 $p=n$, 如果 f_i 之间有线性关系

$$\lambda_1 f_1+\cdots+\lambda_n f_n=0,$$

则有

$$\lambda_1 f_1(x)+\cdots+\lambda_n f_n(x)=0 \quad \text{对于所有 } x\in \mathrm{K}^p.$$

取 x 为 (1) 的一个解, 则有

$$\lambda_1\beta_1+\cdots+\lambda_n\beta_n=0. \tag{7}$$

如果 (1) 对于任意右端常数项都有解, 那么 (7) 对于任意 $\beta_i\in\mathrm{K}$ 都满足, 这显然蕴含

$$\lambda_1=0,\cdots,\lambda_n=0;$$

故 a) 蕴含 c) (同样注意到如果 f_i 不是线性无关的, 方程组 (5) 仅当 β_j 满足非平凡的条件, 即相容性条件时才有解).

剩下的是要证明 c) 蕴含 a). 假定 c) 成立, 根据第 2 小节或 §19 的定理 4, 方程组 (1) 至少总有一个解, 故同态 f 是满射的. 但由于根据假设, K^p 和 K^n 有同样的维数, 同态 f 还是单射的, 因此 (1) 总至多有一个解, 至此证明结束.

称满足定理 1 条件 a), b), c) 的所有线性方程组为 **Cramer 方程组**. 本节的目的本质上在于证明任意线性方程组的求解都可以归结为 Cramer 方程组的求解.

如果事先知道了 $p=n$ (即一个方程组的方程的个数和未知元的个数相同), 可以利用下列定理, 它给出了 Cramer 方程组的更多重要特征.

定理 2 给定一个 n 个未知元 n 个方程的方程组

$$\mathrm{A}x=b,$$

以下性质是等价的:

a) 对于任意 b, 方程组 (4) 有一个且仅一个解.

b) 对于任意 b, 方程组 (4) 至少具有一个解.

c) 对于任意 b, 方程组 (4) 至多具有一个解

d) 存在一个向量 b, 使得方程组 (4) 具有一个且仅一个解.

e) 与 (4) 相伴的齐次方程 $\mathrm{A}y=0$ 仅有平凡解 $y=0$.

f) 矩阵 A 是可逆的.

此外, 如果这些条件满足, 则方程组 (4) 的解

$$x=\mathrm{A}^{-1}b. \tag{8}$$

把 (4) 写成形式 $f(x)=b$, 这里 f 是 K^n 的自同态, 它关于典范基的矩阵是 A, 我们发现 a) 表示 f 是双射, b) 表示 f 是满射, c) 表示 f 是单射, e) 表示 $\mathrm{Ker}(f)=\{0\}$, 而 f) 表示 f 在 K^n 的自同态环内是可逆的 (参见 §15 第 2 小节), 故 a), b), c), e), f) 的等价性是 §19 定理 13 的推论 1 的结果.

此外显然 a) 蕴含 d); 由第 2 小节的推理知道 d) 蕴含 e).

为了完成证明, 只剩下确定公式 (8), 而这是平凡的. 定理证毕.

例 1 假定 $n=2$, 而 K 是交换域. 我们有两个未知元 x 和 y 的两个方程的方程组

$$\begin{cases} ax+by=u, \\ cx+dy=v. \end{cases}$$

根据 §15 第 3 小节, 矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 是可逆的, 当且仅当

$$ad-bc\neq 0,$$

故这个条件刻画了这种情形下 Cramer 方程组的特征. 如果这个条件满足, 容易验证所考虑的方程的解由以下公式给定:

$$x=\frac{du-bv}{ad-bc},\quad y=\frac{av-cu}{ad-bc}.$$

正如在 §24 将要看到的, 行列式理论把这些公式推广到系数在一个交换域 K 内的任意个数方程的 Cramer 方程组. 所有的问题是求计算可逆的方阵的明晰公式.

5. 线性无关的方程组: 化简为 Cramer 方程组

我们要在线性型 $(f_i)_{1\leqslant i\leqslant n}$ 线性无关的情形下考察方程组 (1), 像在第 2 小节看到的那样, 总可以归结为这种情形. 但是不再假设 $p=n$, 这时必定有

$$n\leqslant p.$$

这是由于根据假设 $f_1,\cdots,f_n$ 是 p 维向量空间 (即 K^p 的对偶空间) 的线性无关的元素.

这时矩阵 $\mathrm{A}=(\alpha_{ij})_{1\leqslant i\leqslant n,1\leqslant j\leqslant p}$ 的秩是 n (第 2 小节), 因此 (§19, 定理 16) 可以从 A 抽取 n 阶可逆方阵. 后面我们假定矩阵

$$\mathrm{U}=(\alpha_{ij})_{1\leqslant i,j\leqslant n}$$

是可逆的, 如果有必要调整未知元的编号即可以归结为这种情形.

把方程写成

$$\begin{cases} \alpha_{11}\xi_1+\cdots+\alpha_{1n}\xi_n=\gamma_1, \\ \qquad\cdots\cdots\cdots\cdots \\ \alpha_{n1}\xi_1+\cdots+\alpha_{nn}\xi_n=\gamma_n, \end{cases}\tag{9}$$

其中我们令

$$\gamma_j = \beta_j - (\alpha_{j,n+1}\xi_{n+1} + \cdots + \alpha_{jp}\xi_p). \tag{10}$$

由于 U 是可逆的, (9) 是一个关于未知元 $\xi_1, \cdots, \xi_n$ 的 Cramer 方程组, 因此对于任意右端的值 γ_i 有唯一解: 而 γ_i 仅依赖 $\xi_{n+1}, \cdots, \xi_p$; 因此得到当 $\xi_{n+1}, \cdots, \xi_p$ 事先给定后, 方程组 (1) 有且仅有一个解.

另外, (9) 的解由关系

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \mathrm{U}^{-1} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}$$

给定. 令

$$\mathrm{U}^{-1} = (\nu_{ij})_{1 \leqslant i,j \leqslant n},$$

则有

$$\xi_i = \nu_{i1}\gamma_1 + \cdots + \nu_{in}\gamma_n, \tag{11}$$

将 γ_j 代以它们的值 (10), 显然得到如下公式:

$$\xi_i = \nu_{i1}\beta_1 + \cdots + \nu_{in}\beta_n + \lambda_{i,n+1}\xi_{n+1} + \cdots + \lambda_{ip}\xi_p \quad (1 \leqslant i \leqslant n), \tag{12}$$

其中的 λ_{ik} 是仅依赖方程组 (1) 的系数 α_{ij} 的新的常量.

由于 (1) 等价于 (9) 与 (10) 同时成立, 而 (9) 等价于 (11), 故显然关系 (1) 的集合等价于关系 (12) 的集合. 即有

定理 3 假定线性型 $f_1, \cdots, f_n$ 是线性无关的. 可以对未知元 ξ_i 编号, 使得当未知元 $\xi_{n+1}, \cdots, \xi_p$ 取定任意给定的值时, 方程组 (1) 有且仅有一个解. 这时, 存在仅依赖方程组 (1) 的系数 α_{ij} 的常量 ν_{ij} 和 λ_{ki}, 使得 (1) 的解是所有满足关系 (12) 的序列 $(\xi_1, \cdots, \xi_p)$.

例 2 取三个未知元两个方程的一个方程组

$$\begin{cases} a'x + b'y + c'z = u', \\ a''x + b''y + c''z = u''. \end{cases}$$

方程组的秩是 0, 1 或 2.

如果秩是 0, 则有

$$a' = b' = c' = a'' = b'' = c'' = 0,$$

方程组仅当 $u' = u'' = 0$ 时有解 (这时所有序列 (x, y, z) 都是解).

如果秩是 1, 这就是说系数 $a',\cdots,c''$ 不全是零, 但是出现在给定方程组的左端的线性型是成比例的. 如果 K 是交换的, 这就是说 (§19, 例 4)

$$b'c''-b''c'=a'c''-a''c'=a'b''-a''b'=0. \tag{13}$$

比如假定 $a'\neq 0$, 必然有一个关系

$$a''x+b''y+c''z=\rho(a'x+b'y+c'z),$$

这特别给出 $a''=\rho a'$, 故

$$\rho=a''/a'.$$

在这种情形, 方程组有解当且仅当 $u''=\rho u'$, 即 $u''a'-u'a''=0$. 如果这样, 就归结为求解

$$a'x+b'y+c'z=u'.$$

而由于 $a'\neq 0$, "一般" 解由

$$x=\frac{u'-b'y-c'z}{a}$$

给定, 并且 y 和 z 可以任意选择.

最后, 设方程组的秩是 2, 并且设 K 是交换的, 关系 (13) 不全满足. 比如假定

$$a'c''-a''c'\neq 0,$$

把给定的方程改写为

$$\begin{cases} a'x+c'z=u'-b'y,\\ a''x+c''z=u''-b''y. \end{cases}$$

这就归结为关于 x 和 z 的 Cramer 方程组, 即可以给予 y 任意的值, 那么给定方程组的解由公式

$$\begin{aligned} x&=\frac{c''(u'-b'y)-c'(u''-b''y)}{a'c''-a''c'}=\frac{c''u'-c'u''}{a'c''-a''c'}+\frac{c'b''-c''b'}{a'c''-a''c'}y,\\ z&=\frac{a'u''-a''u'}{a'c''-a''c'}+\frac{a''b'-a'b''}{a'c''-a''c'}y \end{aligned}$$

给出.

§20 习题

用逐次消元法解下列线性方程组[*] (这个方法就是: 利用一个方程通过其他未知元计算一个未知元, 然后把得到的结果代入其余的方程, 用这种方法就得到比初始的方程组少一个方程和一个未知元的方程组).

[*] 在习题 1 至 17 中 (摘自 Proskurjakov 的习题集, 在那里可以找到更多其他习题), 基础域是 **C**. 不过, 愿意在计算中引进更多变形的读者可以在任意交换域 K (并且考虑 K 的特征) 上进行, 或在有理整数环 **Z** 内求解, 只要方程有意义. 自然在研究行列式理论之后, 读者应当把行列式用于这些习题的求解.

1. $\begin{cases} 2x - y + 3z = 9, \\ 3x - 5y + z = -4, \\ 4x - 7y + z = 5. \end{cases}$

2. $\begin{cases} 2x + 3y + 5z = 10, \\ 3x + 7y + 4z = 3, \\ x + 2y + 2z = 3. \end{cases}$

3. $\begin{cases} 5x + 2y + 3z = -2, \\ 2x - 2y + 5z = 0, \\ 3x + 4y + 2z = -10. \end{cases}$

4. $\begin{cases} 4bcx + acy - 2abz = 0, \\ 5bcx + 3acy - 4abz = -abc, \\ 3bcx + 2acy - abz = 4abc \text{ (假定 } abc \neq 0\text{)}. \end{cases}$

5. $\begin{cases} x + y + z = a, \\ x + \omega y + \omega^2 z = b, \\ x + \omega^2 y + \omega z = c \text{ (} \omega \text{ 是 1 的立方根)}. \end{cases}$

6. $\begin{cases} ax - 3y + 5z = 4, \\ x - ay + 3z = 2, \\ 9x - 7y + 8az = 0 \text{ (按照 } a \text{ 的值进行讨论)}. \end{cases}$

7. $\begin{cases} ax + 2z = 2, \\ 5x + 2y = 1, \\ x - 2y + bz = 3 \text{ (按照 } a \text{ 和 } b \text{ 的值进行讨论)}. \end{cases}$

8. $\begin{cases} 2x + 2y - z + t = 4, \\ 4x + 3y - z + 2t = 6, \\ 8x + 5y - 3z + 4t = 12, \\ 3x + 3y - 2z + 2t = 6. \end{cases}$

9. $\begin{cases} 2x - y - 6z + 3t = -1, \\ 7x - 4y + 2z - 15t = -32, \\ x - 2y - 4z + 9t = 5, \\ x - y + 2z - 6t = -8. \end{cases}$

10. $\begin{cases} 2x - 5y + 3z + t = 5, \\ 3x - 7y + 3z - t = -1, \\ 5x - 9y + 6z + 2t = 7, \\ 4x - 6y + 3z + t = 8. \end{cases}$

11. $\begin{cases} 6x + 6y + 5z + 18t + 20u = 14, \\ 10x + 9y + 7z + 24t + 30u = 18, \\ 12x + 12y + 13z + 27t + 35u = 32, \\ 8x + 6y + 6z + 15t + 20u = 16, \\ 4x + 5y + 4z + 15t + 15u = 11. \end{cases}$

12. $\begin{cases} 2x+7y+3z+t=5, \\ x+3y+5z-2t=3, \\ x+5y-9z+8t=1, \\ 5x+18y+4z+5t=12. \end{cases}$

13. $\begin{cases} 2x+5y-8z=8, \\ 4x+3y-9z=9, \\ 2x+3y-5z=7, \\ x+8y-7z=12. \end{cases}$

14. $\begin{cases} 6x+3y+2z+3t+4u=5, \\ 4x+2y+z+2t+3u=4, \\ 4x+2y+3z+2t+u=0, \\ 2x+y+7z+3t+2u=1. \end{cases}$

(此外证明习题 14 的所有解都是整数.)

15. $\begin{cases} 2x+3y+z+2t=3, \\ 4x+6y+3z+4t=5, \\ 6x+9y+5z+6t=7, \\ 8x+12y+7z+\lambda t=9 \text{ (按照 } \lambda \text{ 的值进行讨论)}. \end{cases}$

16. $\begin{cases} ax+y+z=1, \\ x+ay+z=1, \\ x+y+az=1 \text{ (按照 } a \text{ 的值进行讨论)}. \end{cases}$

17. $\begin{cases} (a+1)x+y+z=a^2+3a, \\ x+(a+1)y+z=a^3+3a^2, \\ x+y+(a+1)z=a^4+3a^3 \text{ (按照 } a \text{ 的值进行讨论)}. \end{cases}$

18. 设 K 是一个域. 一方面考虑系数在 K 内的 n 个未知元 n 个线性方程的方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=b_1, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1+\cdots+a_{nn}x_n=b_n, \end{cases} \tag{i}$$

另一方面考虑 $n+1$ 个未知元的齐次线性方程组

$$\begin{cases} a_{11}y_1+\cdots+a_{1n}y_n-b_1y_{n+1}=0, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1+\cdots+a_{nn}y_n-b_ny_{n+1}=0. \end{cases} \tag{ii}$$

a) 证明如果 (ii) 的解满足 $y_{n+1}\neq 0$, 那么

$$x_i=y_iy_{n+1}^{-1}$$

组成 (i) 的解, 并且反之用 (i) 的解可以构造 (ii) 的 $y_{n+1}\neq 0$ 的解.

b) 证明下面两个性质是等价的: (i) 的相伴齐次方程不具有任何非平凡解; (ii) 的所有非平凡解满足 $y_{n+1}\neq 0$.

c) 利用 §19 的定理 7 的推论, 从 b) 推出下列事实的一个证明: §20 的定理 2 的条件 a) 和 e) 是等价的.

19. 如果实系数的线性方程组至少具有一个复数解, 则它具有一个实数解.

¶¶20. 设 L 是一个域, 而 K 是 L 的一个子域.

a) 考虑系数在 K 内的 n 个未知元 n 个齐次线性方程的方程组. 证明如果方程组在 L 内具有一个非平凡解, 则它在 K 内具有一个非平凡解 (比如利用逐次消元法, 关于 n 进行归纳推理).

b) 证明如果矩阵 $\mathrm{A} \in \mathrm{M}_n(\mathrm{K})$ 在环 $\mathrm{M}_n(\mathrm{L})$ 内是可逆的, 则它在环 $\mathrm{M}_n(\mathrm{K})$ 内也是可逆的.

c) 如果 K^n 的元素 (对应的, K^n 上的线性型) 在 K 上是线性无关的, 则它们在 L 上也是线性无关的, 反之亦然.

d) 一个系数和常数项在 K 内的线性方程组在 K 内具有一个解, 必须并且只需它在 L 内具有一个解.

e) 系数在 K 内的齐次线性方程组在 L 内的解是所考虑的方程组在 K 内的解 (系数在 L 内) 的线性组合.

f) 如果系数在 $\mathbf{Z}$ 内的齐次线性方程组在 $\mathbf{C}$ 内具有一个非平凡解, 则它在 $\mathbf{Z}$ 内具有一个非平凡解.

第五章　行列式

正如在第四章的引言中所述，行列式理论主要的目的是提供线性无关性的明晰准则和线性方程组的解的明晰公式.

代替借助 §23 第 5 小节的公式定义行列式和由此公式导出行列式的性质的传统方法，我们陈述行列式理论，要采取“几何的”方法，它以交错多线性型理论为基础，其中包括从行列式的基本性质作为出发点，并从这些性质导出行列式计算的规则. Kronecker 讲授这个理论已经有 80 余年，而且近 15 年来日益推广开来.

第三章展开的模理论在任意的基础环上有效，而行列式理论假定 K 是交换的. 它的一些结果甚至假定 K 是一个域，但我们仅当必需时才做这个假设. 在大多数情形，为了证明所面对的公式，只需要机械计算，并且对于交换环证明它们跟对于域 (比如实数域) 证明它们同样简单 (并且更有用).

初学者觉得 §23 的计算过分困难可以直接跳到 §24 (承认定理 1)，借助习题进行行列式计算的训练. 然后再返回到 §23.

§21　多重线性函数

1. 多重线性映射的定义

设 X, Y 和 M 是一个交换环 K 上的模. 称一个映射

$$f : \mathrm{X} \times \mathrm{Y} \to \mathrm{M}$$

是**双线性的**，如果对于所有 $b \in \mathrm{Y}$, $f(x, b)$ 是 $x \in \mathrm{X}$ 的线性函数，并且对于所有 $a \in \mathrm{X}$,

$f(a,y)$ 是 $y \in \mathrm{Y}$ 的线性函数. 这意味着有等式

$$\begin{aligned} f(x'+x'',y) &= f(x',y)+f(x'',y), \quad f(\lambda x,y)=\lambda f(x,y), \\ f(x,y'+y'') &= f(x,y')+f(x,y''), \quad f(x,\lambda y)=\lambda f(x,y). \end{aligned} \tag{1}$$

现在设 $\mathrm{X},\mathrm{Y},\mathrm{Z}$ 和 M 是 K-模, 称一个映射

$$f: \mathrm{X}\times\mathrm{Y}\times\mathrm{Z}\to\mathrm{M}$$

是**三线性的**, 如果对于给定的 $b\in\mathrm{Y}$ 和 $c\in\mathrm{Z}$, $f(x,b,c)$ 是 $x\in\mathrm{X}$ 的线性函数; 对于给定的 $a\in\mathrm{X}$ 和 $c\in\mathrm{Z}$, $f(a,y,c)$ 是 $y\in\mathrm{Y}$ 的线性函数; 对于给定的 $a\in\mathrm{X}$ 和 $b\in\mathrm{Y}$, $f(a,b,z)$ 是 $z\in\mathrm{Z}$ 的线性函数.

更一般的, 设 $\mathrm{X}_1,\cdots,\mathrm{X}_p$ 和 M 是 K-模, 称形如

$$f:\mathrm{X}_1\times\cdots\times\mathrm{X}_p\to\mathrm{M}$$

的映射是**多重线性的** (更精确的, p-**线性的**), 如果对于所有指标 i, $1\leqslant i\leqslant p$, 和任意的向量

$$a_1\in\mathrm{X}_1,\cdots,a_{i-1}\in\mathrm{X}_{i-1},a_{i+1}\in\mathrm{X}_{i+1},\cdots,a_p\in\mathrm{X}_p,$$

从 X_i 到 M 内的映射

$$x\to f(a_1,\cdots,a_{i-1},x,a_{i+1},\cdots,a_p) \tag{2}$$

是线性的. 即说 f 是多重线性的, 如果给定 $p-1$ 个变量以固定值, 就得到非固定变量的一个线性函数. 可以用推广 (1) 的公式表示这个条件, 即

$$\begin{aligned} &f(x_1,\cdots,x_{i-1},x_i'+x_i'',x_{i+1},\cdots,x_p) \\ =\ &f(x_1,\cdots,x_{i-1},x_i',x_{i+1},\cdots,x_p)+f(x_1,\cdots,x_{i-1},x_i'',x_{i+1},\cdots,x_p), \end{aligned} \tag{3}$$

$$\begin{aligned} &f(x_1,\cdots,x_{i-1},\lambda x_i,x_{i+1},\cdots,x_p) \\ =\ &\lambda f(x_1,\cdots,x_{i-1},x_i,x_{i+1},\cdots,x_p). \end{aligned} \tag{4}$$

用记号

$$\mathscr{L}(\mathrm{X}_1,\cdots,\mathrm{X}_p;\mathrm{M})$$

表示从 $\mathrm{X}_1\times\cdots\times\mathrm{X}_p$ 到 M 内的多重线性映射, 这是从 $\mathrm{X}_1\times\cdots\times\mathrm{X}_p$ 到 M 内的所有映射 (多重线性映射或非多重线性映射) 的模 (§10, 例 4) 的**子模**. 换句话说, 如果 f 和 g 是多重线性的, 则对于所有标量 λ 和 μ, $\lambda f+\mu g$ 也是多重线性的. 这个结果 (直接来自以下事实: 线性映射, 比如形如 (2) 的映射, 的线性组合仍是线性映射) 允许我们认为集合 $\mathscr{L}(\mathrm{X}_1,\cdots,\mathrm{X}_p;\mathrm{M})$ 是 K 上的一个模.

现在给出几个多重线性映射的重要的例子.

例 1 取 $\mathrm{X}_1 = \cdots = \mathrm{X}_p = \mathrm{M} = \mathrm{K}$, 并且令

$$f(x_1, \cdots, x_p) = x_1 \cdots x_p,$$

那么根据交换环的公理 f 是多重线性的. 我们注意到类似于 (1) 或 (3) 的关系以及表达乘法对于加法的分配律的关系成立.

例 2 设 X 和 Y 是两个 K-模, 称所有从 $\mathrm{X} \times \mathrm{Y}$ 到基础环 K 内的双线性映射为**双线性型**. 例如考虑 X 上的线性型 u 和 Y 上的线性型 v, 那么

$$f(x, y) = u(x)v(y)$$

是 $\mathrm{X} \times \mathrm{Y}$ 上的一个双线性型, 因为, 如果比如说给 y 一个固定值 b, 我们得到表达式 $u(x)v(b)$, 这是与 $u(x)$ 成比例的, 从而是 x 的线性型.

更一般的, 如果 $\mathrm{X}_1, \cdots, \mathrm{X}_p$ 是环 K 上的模, 称所有从 $\mathrm{X}_1 \times \cdots \times \mathrm{X}_p$ 到 K 内的多重线性映射为 $\mathrm{X}_1 \times \cdots \times \mathrm{X}_p$ 上的**多重线性型**. 如果对于所有 i 选择 X_i 上的一个线性型 u_i, 并且令

$$f(x_1, \cdots, x_p) = u_1(x_1) \cdots u_p(x_p),$$

就得到 $\mathrm{X}_1 \times \cdots \times \mathrm{X}_p$ 上的一个多重线性型. 称这个多重线性型为线性型 $u_1, \cdots, u_p$ 的**张量积**, 通常用记号

$$f = u_1 \otimes \cdots \otimes u_p$$

表示. 后面 (定理 3) 将证明, 如果 $\mathrm{X}_1, \cdots, \mathrm{X}_p$ 是有限生成自由 K-模, 则 $\mathrm{X}_1 \times \cdots \times \mathrm{X}_p$ 上的所有多重线性型都是线性型的张量积的和.

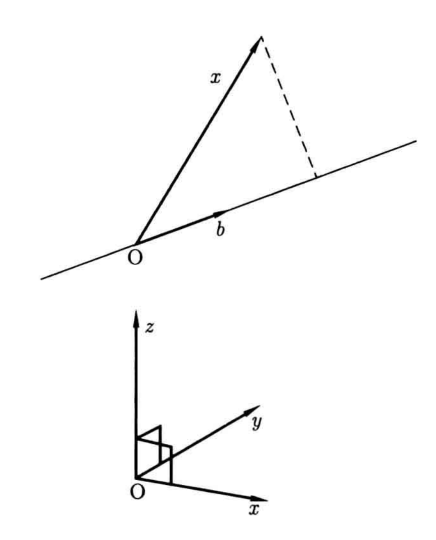
例 3 取 $\mathrm{K} = \mathbf{R}$, 并且用 E 表示通常三维空间内起点为 O 的向量组成的向量空间. 假定选择了一个单位长度, 可以定义两个向量 x 和 y **的标量积**: 这是 x 和 y 的长度以及 x 到 y 的角的余弦相乘得到的数. 用下列记号之一表示标量积:

$$x \cdot y, \quad (x, y), \quad (x|y).$$

我们将只用第三个记号 (第二个已经用来表示 x 和 y 组成的序偶). 标量积有了定义, 那么表达式 $(x|y)$ 是 $\mathrm{E} \times \mathrm{E}$ 上的一个双线性型. 比如为了证明 $(x|b)$ 是 x 的线性函数, 只需注意这个数等于向量 x 在向量 b 所在的直线上的正交投影的 (有向) 长度; 而在一条固定直线上的投影这一运算是线性的[*].

---

[*] 留给读者仔细展开几何考虑的细节, 这些考虑是本例和下例的基础.



例 4　K 和 E 跟前一个例一样, 考虑两个向量 x 和 y 的向量积: 这是一个向量 z, 其长度是 x 和 y 的长度以及它们的夹角的正弦的乘积, 即以 x 和 y 为边的平行四边形的面积, 它正交于由 x 和 y 定义的平面, 并且它的方向使得标架 xyz 是正向的 (见图), (正向的标架的概念不是数学意义上的, 唯一正确的概念是两个标架同向的概念, 在 §23 注 4 将会看到这可以通过行列式定义). 把它记作

$$x \times y \quad 或 \quad x \wedge y;$$

我们将仅采用第二个记号. 有了这个定义, 容易发现 (读者应当作为习题证明之) 从 $\mathrm{E} \times \mathrm{E}$ 到 E 内的映射

$$(x, y) \to x \wedge y$$

是双线性的.

我们要注意以下交换性公式:

$$(x|y) = (y|x); \quad x \wedge y = -y \wedge x.$$

例 5　K 和 E 仍然和前面一样, 称数

$$(x|y|z) = (x|y \wedge z)$$

为三个向量 x, y, z 的**混合积**, 这是向量 x 与 y 和 z 的向量积的标量积. 混合积是 x, y 和 z 的三线性函数, $(x|y|z)$ 的绝对值等于在向量 x, y, z 上所构建的平行六面体的体积, 而 $(x|y|z)$ 的符号是正的, 如果标架 xyz 是正向的; 否则是负的, 如果标架 xyz 是反向的. 由此我们有关系

$$(x|y|z) = (y|z|x) = (z|x|y) = -(x|z|y) = -(y|x|z) = -(z|y|x).$$

¶例 6 设 K 是任意一个交换环, X 是一个 K-模, 而 X^* 是 X 的对偶模 (§16, 第 1 小节). 给定了整数 $p, q \geqslant 0$, 称所有从

$$(X^*)^p \times X^q = \underbrace{X^* \times \cdots \times X^*}_{p} \times \underbrace{X \times \cdots \times X}_{q}$$

到基础环 K 内的 $(p+q)$-线性映射为 p **次共变** q **次反变的张量**, 或 $\begin{pmatrix} p \\ q \end{pmatrix}$ **型张量**. 这样的张量是一个在 K 内取值的函数 $f(u_1, \cdots, u_p, x_1, \cdots, x_q)$, 当 $u_1, \cdots, u_p$ 是 X 上的线性型且 $x_1, \cdots, x_q$ 是 X 的向量时有定义, 并且线性地依赖每一个变量 $u_1, \cdots, x_q$. 在特殊情形, 我们称所有 $\begin{pmatrix} 0 \\ q \end{pmatrix}$ 型张量为 X 上的 q-线性型, 这就是从 X^q 到 K 内的多重线性型.

在 X 上的张量中出现 X 上的线性型: 这就是 $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 型的张量. 此外, X 的每个向量定义 X 上的一个 $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 型张量, 即 X^* 上的一个线性型, 在 §16, 第 3 小节引进了线性型 $u \to u(x)$, 它把一个模嵌入到它的对偶中. 当 X 是**有限生成**的和自由的时候, 在 X 和 $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 型的张量之间的对应是双射的 (§16, 定理 2), 这就使我们能够把这两个概念视作等同, 尤其在物理学中正是这样做的.

大部分物理学家给张量 (在 $\mathbf{R}$ 上的小维数的向量空间的情形下) 一个明显比前面我们所给的更复杂的定义 (第 5 小节, 注 3). 此外, 在物理学中, 人们仅对张量场感兴趣. 所谓**张量场**就是从所考虑的向量空间 X 到 X 上所有的给定型的张量的集合内的一个映射, 即 X 上的一个函数, 它在 X 的每个点的值是 X 上的给定型的张量. 张量在物理学中的应用长期以来难以理解, 其原因是没有区分张量 (在这个词的这里所给的意义下) 和取值为张量的函数; 这与下列状况类似, 谈论向量值函数却没有事先定义什么是向量. 对物理学家说句公道话, 我们应该说, 向量空间上的线性型的概念 (没有这个概念实际上不可能给张量以简单的定义) 从 1930 年以后才开始在数学里推广开来.

2. 多重线性映射的张量积

设

$$f : X_1 \times \cdots \times X_p \to K, \quad g : Y_1 \times \cdots \times Y_q \to K$$

是多重线性型. 称由

$$f \otimes g(x_1, \cdots, x_p, y_1, \cdots, y_q) = f(x_1, \cdots, x_p) g(y_1, \cdots, y_q) \tag{5}$$

定义的映射

$$f \otimes g : \mathrm{X}_1 \times \cdots \times \mathrm{X}_p \times \mathrm{Y}_1 \times \cdots \times \mathrm{Y}_q \to \mathrm{K}$$

为 f **和** g **的张量积**. 这仍然是一个多重线性映射, 因为如果比如说给 $x_2, \cdots, y_q$ 以固定值 $a_2, \cdots, b_q$, 这样得到的表达式

$$f(x_1, a_2, \cdots, a_p)g(b_1, \cdots, b_q)$$

作为 x_1 的函数正比于 $f(x_1, a_2, \cdots, a_p)$, 因此跟后者一样是 x_1 的线性函数.

如果有三个多重线性型 f, g, h, 我们有结合性关系

$$(f \otimes g) \otimes h = f \otimes (g \otimes h),$$

并且两端的公共值正如我们立刻看到的是

$$f(x_1, \cdots, x_p)g(y_1, \cdots, y_q)h(z_1, \cdots, z_r).$$

这就让我们可以定义任意多个多重线性型的张量积, 并且推广例 2 对线性型所引进的概念.

我们注意公式

$$f \otimes g = g \otimes f$$

即使两端有同样的型也是错误的. 比如如果 f 和 g 是模 X 上的线性型, 则有

$$(f \otimes g)(x, y) = f(x)g(y)$$
$$(g \otimes f)(x, y) = g(x)f(y),$$

这些表达式没有任何理由是相等的.

¶例 7 考虑模 X 上的一个 $\begin{pmatrix} p \\ q \end{pmatrix}$ 型的张量和一个 $\begin{pmatrix} r \\ s \end{pmatrix}$ 型张量, 那么 $f \otimes g$ 是在笛卡儿乘积空间

$$(\mathrm{X}^*)^p \times \mathrm{X}^q \times (\mathrm{X}^*)^r \times \mathrm{X}^s$$

上的多重线性型. 一般把 $(\mathrm{X}^*)^p \times \mathrm{X}^q \times (\mathrm{X}^*)^r \times \mathrm{X}^s$ 与

$$(\mathrm{X}^*)^{p+r} \times \mathrm{X}^{q+s}$$

等同, 这就允许把 $f \otimes g$ 看作 X 上的 $\begin{pmatrix} p+r \\ q+s \end{pmatrix}$ 型的一个张量, 其定义是对于任意 $u_i \in \mathrm{X}^*$ 和 $x_j \in \mathrm{X}$,

$$\begin{aligned} &(f \otimes g)(u_1, \cdots, u_{p+r}, x_1, \cdots, x_{q+s}) \\ &= f(u_1, \cdots, u_p, x_1, \cdots, x_q)g(u_{p+1}, \cdots, u_{p+r}, x_{q+1}, \cdots, x_{q+s}). \end{aligned}$$

我们说 $f\otimes g$ 是张量 f 和 g 的张量积.

例如考虑两个向量 $a,b\in \mathrm{X}$ 和三个线性型 $f,g,h\in \mathrm{X}^*$. 把 a 和 b 等同于张量(例 6), 可以定义张量 $a\otimes b\otimes f\otimes g\otimes h=\varphi$, 它是二次共变和三次反变的, 即这是两个线性型 $u,v\in \mathrm{X}^*$ 和三个向量 $x,y,z\in \mathrm{X}$ 的多重线性型. 容易验证事实上有

$$\varphi(u,v,x,y,z)=u(a)v(b)f(x)g(y)h(z).$$

注 1 在定义 $f\otimes g$ 的公式 (5) 中, 如果想获得一个多重线性函数, 本质的一点是变量 y_j 独立于变量 x_i. 比如说, 如果 f 和 g 是模 L 上的两个线性型, 表达式 $f(x)g(y)$ 是 $\mathrm{L}\times\mathrm{L}$ 上的双线性型, 但是函数 $f(x)g(x)$ 一般则不是 L 上的线性型.

3. 几个代数等式

当实施一个环内的代数计算时, 经常要计算一个乘积, 它的每个因子是一些项的和, 并且要借助下列规则展开它: 为了求几个和的乘积, 在每个和里任意取一项, 把取出的项相乘, 再把所得的各个乘积相加. 这个规则可以翻译成公式: 例如, 如果 $(x_i)_{i\in\mathrm{I}}$ 和 $(y_j)_{j\in\mathrm{J}}$ 是环 K 的元素的两个有限族, 则有

$$\sum_{i\in\mathrm{I}}x_i\cdot\sum_{j\in\mathrm{J}}y_j=\sum_{i\in\mathrm{I},j\in\mathrm{J}}x_iy_j. \tag{6}$$

给定一个环的元素的三个有限族 $(x_i)_{i\in\mathrm{I}}$, $(y_j)_{j\in\mathrm{J}}$ 和 $(z_k)_{k\in\mathrm{K}}$, 则有关系

$$\sum_{i\in\mathrm{I}}x_i\cdot\sum_{j\in\mathrm{J}}y_j\cdot\sum_{k\in\mathrm{K}}z_k=\sum_{i\in\mathrm{I},j\in\mathrm{J},k\in\mathrm{K}}x_iy_jz_k. \tag{7}$$

更一般的, 假定给定一个环的元素的 p 个有限族, 把这些族记作 $(x_{1i_1})_{i_1\in\mathrm{I}_1},\cdots,(x_{pi_p})_{i_p\in\mathrm{I}_p}$, 则有

$$\sum_{i_1\in\mathrm{I}_1}x_{1i_1}\cdots\sum_{i_p\in\mathrm{I}_p}x_{pi_p}=\sum_{i_1\in\mathrm{I}_1,\cdots i_p\in\mathrm{I}_p}x_{1i_1}\cdots x_{pi_p}. \tag{8}$$



注 2 在公式 (6) 里经常会遇到 $\mathrm{I}=\mathrm{J}$ 的情形, 这时给定的族记成 $(x_i)_{i\in\mathrm{I}}$ 和 $(y_j)_{j\in\mathrm{I}}$. 但是要注意在这种情形下, 不能把关系 (6) 写成

$$\sum_{i\in\mathrm{I}}x_i\cdot\sum_{i\in\mathrm{I}}y_i=\sum_{i\in\mathrm{I}}x_iy_i,$$

因为这个关系显然是错误的. 例如, 相信总有

$$(x'+x'')(y'+y'')=x'y'+x''y''$$

是毫无道理的. 避免这类错误的有效方法是出现在两个不同的和里的指标绝不用同一个字母表示. 还要记住求和指标事实上并不出现在求和的结果中, 仅起

表示要执行的操作的缩写的作用, 总可以用未曾使用的其他字母代替它. 例如, 我们有加法群的元素 $x_1,\cdots,x_n$, 表达式

$$\sum_{1\leqslant i\leqslant n} x_i,\quad \sum_{1\leqslant h\leqslant n} x_h,\quad \sum_{1\leqslant \lambda\leqslant n} x_\lambda$$

是相同的.

对于多重线性映射可以证明类似于 (6), (7) 和 (8) 的公式, 理由在于蕴含 (6), (7) 和 (8) 的 "乘法对于加法的分配律" 根据定义在多重线性映射的情形也是满足的.

例如考虑一个交换环 K 上的模 X, Y, M 和从 X × Y 到 M 内的一个双线性映射 f. 给定 X 的元素的一个有限族 $(x_i)_{i\in \mathrm{I}}$ 和 Y 的元素的一个有限族 $(y_j)_{j\in \mathrm{J}}$, 则有关系

$$f\left(\sum_{i\in \mathrm{I}} x_i, \sum_{j\in \mathrm{J}} y_j\right) = \sum_{i\in \mathrm{I}, j\in \mathrm{J}} f(x_i, y_j). \tag{6'}$$

为了确认这个等式, 令

$$a = \sum_{i\in \mathrm{I}} x_i,\quad f_a(y) = f(a, y).$$

由于 f 是双线性的, f_a 是从 Y 到 M 内的线性映射, 因此有

$$f\left(a, \sum_{j\in \mathrm{J}} y_j\right) = f_a\left(\sum_{j\in \mathrm{J}} y_j\right) = \sum_{j\in \mathrm{J}} f_a(y_j) = \sum_{j\in \mathrm{J}} f(a, y_j);$$

再令 $f_j(x) = f(x, y_j)$, 基于同样的理由, 它是一个从 X 到 M 内的线性映射, 因此有

$$f(a, y_j) = f_j(a) = f_j\left(\sum_{i\in \mathrm{I}} x_i\right) = \sum_{i\in \mathrm{I}} f_j(x_i) = \sum_{i\in \mathrm{I}} f(x_i, y_j),$$

把这个结果代入前式显然得到 (6′).

同样设 X, Y, Z 和 M 是 K-模, 而 f 是从 X × Y × Z 到 M 内的一个三重线性映射. 如果 $(x_i)_{i\in \mathrm{I}}$, $(y_j)_{j\in \mathrm{J}}$ 和 $(z_k)_{k\in \mathrm{K}}$ 分别是 X, Y 和 Z 的元素的有限族, 则有

$$f\left(\sum_{i\in \mathrm{I}} x_i, \sum_{j\in \mathrm{J}} y_j, \sum_{k\in \mathrm{K}} z_k\right) = \sum_{i\in \mathrm{I}, j\in \mathrm{J}, k\in \mathrm{K}} f(x_i, y_j, z_k). \tag{7'}$$

事实上, 令

$$c = \sum_{k\in \mathrm{K}} z_k,\quad f_c(x, y) = f(x, y, c).$$

显然 f_c 是双线性的, 对于 f_c 应用 (6′) 即得 (7′) 的左端等于

$$\sum_{i\in \mathrm{I}, j\in \mathrm{J}} f_c(x_i, y_j) = \sum_{i\in \mathrm{I}, j\in \mathrm{J}} f\left(x_i, y_j, \sum_{k\in \mathrm{K}} z_k\right),$$

但是从 Z 到 M 内的映射

$$f_{ij}(z) = f(x_i, y_j, z)$$

是线性的, 刚写出的和的一般项是 f_{ij} 在 $\sum_{k\in K} z_k$ 的值, 故这一项等于

$$\sum_{k\in K} f_{ij}(z_k) = \sum_{k\in K} f(x_i, y_j, z_k).$$

终于得到 (7′) 的左端等于

$$\sum_{i\in I, j\in J} \sum_{k\in K} f(x_i, y_j, z_k) = \sum_{i\in I, j\in J, k\in K} f(x_i, y_j, z_k),$$

这就证明了 (7′).

最后为了推广 (8), 考虑 p-重线性映射

$$f : X_1 \times \cdots \times X_p \to M,$$

并且在每个模 X_h 里选择向量的有限族 $(x_{hi_h})_{i_h\in I_h}$, 则有等式

$$f\left(\sum_{i_1\in I_1} x_{1i_1}, \cdots, \sum_{i_p\in I_p} x_{pi_p}\right) = \sum_{i_1\in I_1, \cdots, i_p\in I_p} f(x_{1i_1}, \cdots, x_{pi_p}). \tag{8′}$$

这个等式通过关于 p 的归纳法证明. 令

$$c = \sum_{i_p\in I_p} x_{pi_p}, \quad f_c(x_1, \cdots, x_{p-1}) = f(x_1, \cdots, x_{p-1}, c),$$

就得到一个 $(p-1)$-重线性映射 f_c. 由归纳假设给出

$$\begin{aligned} f_c\left(\sum_{i_1\in I_1} x_{1i_1}, \cdots, \sum_{i_{p-1}\in I_{p-1}} x_{p-1,i_{p-1}}\right) &= \sum_{i_1\in I_1, \cdots, i_{p-1}\in I_{p-1}} f_c(x_{1i_1}, \cdots, x_{p-1,i_{p-1}}) \\ &= \sum_{i_1\in I_1, \cdots, i_{p-1}\in I_{p-1}} f_{i_1, \cdots, i_{p-1}}\left(\sum_{i_{p-1}\in I_p} x_{pi_p}\right), \end{aligned}$$

其中

$$f_{i_1\cdots i_{p-1}}(x) = f(x_{1i_1}, \cdots, x_{p-1,i_{p-1}}, x) \quad 对于\ x \in X_p.$$

而这个表达式是 x 的线性函数, 故我们发现 (8′) 的左端等于

$$\sum_{i_1\in I_1, \cdots, i_{p-1}\in I_{p-1}} \sum_{i_p\in I_p} f_{i_1\cdots i_{p-1}}(x_{pi_p}),$$

由于

$$f_{i_1\cdots i_{p-1}}(x_{pi_p}) = f(x_{1i_1}, \cdots, x_{pi_p}),$$

我们得到 (8′).

我们注意到在这些公式中从没有涉及基础环 K, 即仅用到第 1 小节的等式 (3), 而没有用到等式 (4). 为了在 (8′) 中涉及 (4), 再选择 K 的元素的族

$$(\lambda_{1i_1})_{i_1\in \mathrm{I}_1}, \cdots, (\lambda_{pi_p})_{i_p\in \mathrm{I}_p},$$

在 (8′) 里把每个 x_{hi_h} 代换成 $\lambda_{hi_h}x_{hi_h}$, 那么 (8′) 右端的一般项则换成

$$\lambda_{1i_1}\cdots\lambda_{pi_p}f(x_{1i_1}, \cdots, x_{pi_p}),$$

理由是有一般的公式

$$f(\xi_1 x_1, \cdots, \xi_p x_p) = \xi_1\cdots\xi_p f(x_1, \cdots, x_p). \tag{9}$$

有了这些铺垫, 我们发现 (8) 就转变为更一般的公式

$$f\left(\sum_{i_1\in \mathrm{I}_1}\lambda_{1i_1}x_{1i_1}, \cdots, \sum_{i_p\in \mathrm{I}_p}\lambda_{pi_p}x_{pi_p}\right) = \sum_{i_1\in \mathrm{I}_1, \cdots, i_p\in \mathrm{I}_p}\lambda_{1i_1}\cdots\lambda_{pi_p}f(x_{1i_1}, \cdots, x_{pi_p}). \tag{10}$$

这个公式使我们能够计算 f 在向量的线性组合上的值. 对于 $p = 2$, 这个公式写成

$$f\left(\sum_{i\in \mathrm{I}}\lambda_i x_i, \sum_{j\in \mathrm{J}}\mu_j y_j\right) = \sum_{i\in \mathrm{I}, j\in \mathrm{J}}\lambda_i\mu_j f(x_i, y_j), \tag{11}$$

而对于 $p = 3$, 则得到

$$f\left(\sum_{i\in \mathrm{I}}\lambda_i x_i, \sum_{j\in \mathrm{J}}\mu_j y_j, \sum_{k\in \mathrm{K}}\nu_k z_k\right) = \sum_{i\in \mathrm{I}, j\in \mathrm{J}, k\in \mathrm{K}}\lambda_i\mu_j\nu_k f(x_i y_j z_k). \tag{12}$$

4. 有限生成自由模的情形

当模 $\mathrm{X}_1, \cdots, \mathrm{X}_p$ 是有限生成的自由的时候 (例如当它们是一个域上的有限维向量空间的时候, 长期以来的最重要的情形), 前一小节的公式让我们能够确定所有的多重线性映射.

首先考察 $p = 2$ 这种最简单的情形.

定理 1 设 X, Y 和 M 是交换环 K 上的模. 假定 X 和 Y 是有限生成自由的, 而 $(a_i)_{1\leqslant i\leqslant m}$ 是 X 的一个基, $(b_j)_{1\leqslant j\leqslant n}$ 是 Y 的一个基. 从 $\mathrm{X}\times\mathrm{Y}$ 到 M 内的一个映射 f 是双线性的, 必须并且只需存在 $c_{ij}\in \mathrm{M}$, 使得对于任意

$$x = \sum_{1\leqslant i\leqslant m}\xi_i a_i \in \mathrm{X} \quad 和 \quad y = \sum_{1\leqslant j\leqslant n}\eta_i b_i \in \mathrm{Y}$$

有

$$f(x,y)=\sum_{1\leqslant i\leqslant m,1\leqslant j\leqslant n}\xi_i\eta_j c_{ij}. \tag{13}$$

当这个条件满足时必然有

$$c_{ij}=f(a_i,b_j). \tag{14}$$

公式 (11) 表明, 如果 f 是双线性的, 必然有

$$f(x,y)=\sum_{1\leqslant i\leqslant m,1\leqslant j\leqslant n}\xi_i\eta_j f(a_i,b_j),$$

于是 f 必然由形式 (13) 这种类型的公式给定. 反之, 假定 f 由 (13) 给定, 为了证明 f 是双线性的, 只需证明和式 (13) 的一般项 $\xi_i\eta_j c_{ij}$ 是 x 和 y 的双线性函数. 由于 c_{ij} 不依赖 x 和 y, 只需证明 $\xi_i\eta_j$ 是双线性的. 如果用 u_i 表示 X 关于基 (a_i) 的坐标函数, 用 v_j 表示 Y 关于基 (b_j) 的坐标函数, 则有

$$\xi_i\eta_j=u_i(x)v_j(y),$$

这就证明了这个表达式是 x 和 y 的双线性函数 (例 2).

还需要证明的是关系 (13), 即

$$f(x,y)=\sum_{1\leqslant i\leqslant m,1\leqslant j\leqslant n}u_i(x)v_j(y)c_{ij}$$

必然蕴含 $c_{ij}=f(a_i,b_j)$. 我们有

$$f(a_i,b_j)=\sum_{k,h}u_k(a_i)v_h(b_j)c_{kh},$$

而

$$u_k(a_i)=\begin{cases}0, & k\neq i,\\ 1, & k=i,\end{cases}\qquad v_h(b_j)=\begin{cases}0, & h\neq j,\\ 1, & h=j.\end{cases}$$

(参见 §16, 第 2 小节, 或直接观察

$$a_i=0\cdot a_1+\cdots+0\cdot a_{i-1}+1\cdot a_i+0\cdot a_{i+1}+\cdots+0\cdot a_m.)$$

定义 $f(a_i,b_j)$ 的和式中仅有的可能的非零项是当 $k=i$ 和 $h=j$ 时的项, 此项缩减为 c_{ij}, 这就完成了定理的证明.

M 的元素 c_{ij} 称为 f 关于 X 的基 (a_i) 和 Y 的基 (b_j) 的**系数**, 当 X = Y 时, 习惯上在 X 和 Y 内取同样的基 (a_i), 这时 $c_{ij}=f(a_i,a_j)$ 称为 f 关于 X 的这个基 (a_i) 的系数.

当 M = K 时通常把 (13) 写成形式

$$f(x,y)=\sum\gamma_{ij}\xi_i\eta_j,\quad \text{其中 } \gamma_{ij}=f(a_i,b_j)\in \mathrm{K}.$$

由于

$$\xi_i\eta_j = u_i(x) \otimes v_j(y),$$

前面的公式还可以在 $\mathrm{X} \times \mathrm{X}$ 上的双线性型的模 $\mathscr{L}(\mathrm{X}, \mathrm{Y}; \mathrm{K})$ 中写成形式

$$f = \sum \gamma_{ij} u_i \otimes v_j,$$

由于这个分解是唯一的, 因此就推导出 mn 个形式 $u_i \otimes v_j$ 构成 $\mathscr{L}(\mathrm{X}, \mathrm{Y}; \mathrm{K})$ 的基.

例 8 考虑例 3 的双线性型. 给定向量空间 E 的一个基 a_1, a_2, a_3, 则有

$$(x|y) = \sum (a_i|a_j) \cdot \xi_i\eta_j.$$

当基 a_1, a_2, a_3 是**正交规范基**时这个公式可以化简, 这时基由长度为 1 并且两两正交的向量组成, 即在直角坐标系下进行计算, 故有

$$(a_i|a_j) = \begin{cases} 0, & i \neq j, \\ 1, & i = j, \end{cases}$$

因此有

$$(x|y) = \xi_1\eta_1 + \xi_2\eta_2 + \xi_3\eta_3,$$

用这个公式计算直角坐标系下两个向量的标量积.

这个公式还可以用来计算直角坐标系下的空间中两个点 P 和 Q 之间的距离. 这个距离其实就是向量

$$\overrightarrow{\mathrm{PQ}} = \overrightarrow{\mathrm{OQ}} - \overrightarrow{\mathrm{OP}}$$

的长度. 设 P 的坐标是 (x, y, z), Q 的坐标是 (x', y', z'), 向量 $\overrightarrow{\mathrm{PQ}}$ (或准确地说, 等于 $\overrightarrow{\mathrm{PQ}}$ 的起点在 O 的向量) 的坐标是 $(x'-x, y'-y, z'-z)$; 它的长度即它同自己的标量积的平方根

$$\sqrt{(x'-x)^2 + (y'-y)^2 + (z'-z)^2},$$

这就是在直角坐标系下的点 P 和 Q 之间的距离的表达式.

例 9 在前一个例子中用例 4 的向量积 $x \wedge y$ 代替标量积 $(x|y)$. 我们有

$$x \wedge y = \sum \xi_i\eta_j a_i \wedge a_j,$$

对于任意选择的基都有

$$a_i \wedge a_i = 0, \quad a_i \wedge a_j = -a_j \wedge a_i;$$

于是有公式

$$x \wedge y = (\xi_2\eta_3 - \xi_3\eta_2)a_2 \wedge a_3 + (\xi_3\eta_1 - \xi_1\eta_3)a_3 \wedge a_1 + (\xi_1\eta_2 - \xi_2\eta_1)a_1 \wedge a_2;$$

如果基是正交规范的, 则有

$$a_2 \wedge a_3 = a_1, \quad a_3 \wedge a_1 = a_2, \quad a_1 \wedge a_2 = a_3,$$

我们就得到在直角坐标系下的公式

$$x \wedge y = (\xi_2\eta_3 - \xi_3\eta_2)a_1 + (\xi_3\eta_1 - \xi_1\eta_3)a_2 + (\xi_1\eta_2 - \xi_2\eta_1)a_3.$$

这些公式以及前一个例子中的公式在实际应用中 (三维解析几何、力学、物理学等) 是极其基本的, 并且依据的只不过就是标量积和向量积的双线性.

对于三重线性映射有一个类似于定理 1 的结果:

定理 2 设 X, Y, Z 和 M 是交换环 K 上的模. 假定 X, Y 和 Z 是有限生成自由的, 而 $(a_i)_{1\leqslant i\leqslant m}$ 是 X 的一个基, $(b_j)_{1\leqslant j\leqslant n}$ 是 Y 的一个基, $(c_k)_{1\leqslant k\leqslant p}$ 是 Z 的一个基. 从 X × Y × Z 到 M 内的一个映射 f 是三线性的, 必须并且只需存在 $c_{ijk} \in \mathrm{M}$, 使得对于任意

$$x = \sum_{1\leqslant i\leqslant m} \xi_i a_i \in \mathrm{X}, \quad y = \sum_{1\leqslant j\leqslant n} \eta_i b_i \in \mathrm{Y}, \quad z = \sum_{1\leqslant k\leqslant p} \zeta_k c_k \in \mathrm{Z}$$

有

$$f(x,y,z) = \sum_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n, 1\leqslant k\leqslant p} \xi_i\eta_j\zeta_k c_{ijk}. \tag{15}$$

当这个条件满足时必然有

$$c_{ijk} = f(a_i, b_j, c_k). \tag{16}$$

公式 (12) 表明如果 f 是三线性的, 必然有关系 (15), 其中的 c_{ijk} 由 (16) 给定. 反之, 假定 f 由关系 (15) 给定, 为了证明 f 是三线性的, 只需证明函数 $\xi_i\eta_j\zeta_k c_{ijk}$ 是三线性的, 由于 c_{ijk} 不依赖 x, y, z, 甚至只需证明 $\xi_i\eta_j\zeta_k$ 是三线性的. 如果引进模 X, Y, Z 关于所考虑的基的坐标函数 u_i, v_j, w_k, 则显然有

$$\xi_i\eta_j\zeta_k = u_i(x)v_j(y)w_k(z),$$

这个表达式的右端必然是 X × Y × Z 上的三重线性函数, 即

$$u_i \otimes v_j \otimes w_k.$$

为了完成证明, 还需要证明关系 (15) 蕴含 (16), 而从 (15) 得到

$$f(a_i, b_j, c_j) = \sum_{\lambda,\mu,\nu} u_\lambda(a_i)v_\mu(b_j)w_\nu(c_k)c_{\lambda\mu\nu}.$$

右端仅有的可能的非零项是当 $\lambda = i$, $\mu = j$ 和 $\nu = k$ 时的项, 此时 $u_\lambda(a_i) = v_\mu(b_j) = w_\nu(c_k) = 1$, 故得 (16).

这里 M 的元素 c_{ijk} 还是称为 f 关于 X, Y, Z 的基 (a_i), (b_j) 和 (c_k) 的**系数**. 当 $X = Y = Z$ 时, 一般在 X, Y 和 Z 中使用同样的基 (a_i), 这时 $c_{ijk} = f(a_i, a_j, a_k)$ 是 f 关于 X 的基 (a_i) 的系数.

当 $M = K$ 时通常把 (15) 写成形式

$$f(x, y, z) = \sum \gamma_{ijk}\xi_i\eta_j\zeta_k, \quad \text{其中 } \gamma_{ijk} = f(a_i, b_j, c_k).$$

这时在 $X \times Y \times Z$ 上的三重线性型的模 $\mathscr{L}(X, Y, Z; K)$ 中有

$$f = \sum \gamma_{ijk} u_i \otimes v_j \otimes w_k,$$

由于这个分解是唯一的, 这就推导出 mnp 个形式 $u_i \otimes v_j \otimes w_k$ 构成 $\mathscr{L}(X, Y, Z; K)$ 的一个基.

例 10 考虑例 5 的混合积 $(x|y|z)$. 我们有

$$(x|y|z) = \sum (a_i|a_j|a_k)\xi_i\eta_j\zeta_k,$$

而如果指标不是两两不同, 则有 $(a_i|a_j|a_k) = 0$, 并且

$$\begin{aligned}(a_1|a_2|a_3) &= (a_2|a_3|a_1) = (a_3|a_1|a_2) = -(a_1|a_3|a_2)\\ &= -(a_2|a_1|a_3) = -(a_3|a_2|a_1),\end{aligned}$$

因此有

$$(x|y|z) = (a_1|a_2|a_3) \cdot (\xi_1\eta_2\zeta_3 + \xi_2\eta_3\zeta_1 + \xi_3\eta_1\zeta_2 - \xi_1\eta_3\zeta_2 - \xi_2\eta_1\zeta_3 - \xi_3\eta_2\zeta_1),$$

如果基 a_1, a_2, a_3 是正交规范的, 则有

$$(x|y|z) = (\xi_1\eta_2\zeta_3 + \xi_2\eta_3\zeta_1 + \xi_3\eta_1\zeta_2 - \xi_1\eta_3\zeta_2 - \xi_2\eta_1\zeta_3 - \xi_3\eta_2\zeta_1).$$

习惯上这个关系的右端用紧凑的记号

$$\begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix}$$

表示. 下一节还要回到这个表达式.

¶**例 11** 设 T 是一个模 X 上的一个两次共变和一次反变的张量, 即在

$$X^* \times X^* \times X$$

上的三重线性型. 假定 X 具有一个基 $(a_1, \cdots, a_n)$, 并且用记号(*) $(a^1, \cdots, a^n)$ 表示它在对偶模 X* (§16, 第 2 小节) 内的对偶基, 换句话说, 线性型 a^i 正是模 X 关于基 (a_i) 的坐标函数. 为了对于 $u, v \in \mathrm{X}^*$ 和 $x \in \mathrm{X}$ 计算 $\mathrm{T}(u, v, x)$, 令

$$\begin{aligned} u &= \sum \alpha_i a^i \qquad \text{由此得} \quad \alpha_i = u(a_i), \\ v &= \sum \beta_i a^i \qquad \text{由此得} \quad \beta_i = v(a_i), \\ x &= \sum \xi^i \cdot a_i, \end{aligned}$$

那么就有

$$\mathrm{T}(u, v, x) = \sum \mathrm{T}_k^{ij} \alpha_i \beta_j \xi^k,$$

其中的常量是

$$\mathrm{T}_k^{ij} = \mathrm{T}(a^i, a^j, a_k),$$

称为张量关于 X 的基 (a_i) 的**系数** (或**分量**, 或**坐标**).

作为结束, 我们要推广定理 1 和 2 到 p-重线性映射, 这里 p 是任意整数.

定理 3 设 $\mathrm{X}_1, \cdots, \mathrm{X}_p$ 和 M 是交换环 K 上的模. 假定 $\mathrm{X}_1, \cdots, \mathrm{X}_p$ 是有限生成的自由的, 并且对于所有指标 $h, 1 \leqslant h \leqslant p$, 设 $(a_{h1}, a_{h2}, \cdots, a_{hn_h})$ 是 X_h 的一个基. 从 $\mathrm{X}_1 \times \cdots \times \mathrm{X}_p$ 到 M 内的一个映射 f 是多重线性的, 必须并且只需存在常向量

$$c_{i_1 \cdots i_p} \in \mathrm{M} \ (1 \leqslant i_1 \leqslant n_1, \cdots, 1 \leqslant i_p \leqslant n_p),$$

使得对于任意

$$x_h = \sum_{1 \leqslant i_h \leqslant n_h} \xi_{h i_h} a_{h i_h}$$

有

$$f(x_1, \cdots, x_p) = \sum_{1 \leqslant i_1 \leqslant n_1, \cdots, 1 \leqslant i_p \leqslant n_p} \xi_{1 i_1} \cdots \xi_{p i_p} c_{i_1 \cdots i_p}. \tag{17}$$

如果上述条件满足, 必定有

$$c_{i_1 \cdots i_p} = f(a_{1 i_1}, \cdots, a_{p i_p}). \tag{18}$$

当 f 是多重线性映射时, 公式 (17) 从第 3 小节的关系 (10) 直接得到. 反之, 为了证明 (17) 总表示多重线性映射, 只需证明

$$u_{i_1 \cdots i_p}(x_1, \cdots, x_p) = \xi_{1 i_1} \cdots \xi_{p i_p}$$

(*) 后面公式中的上指标自然不是指数. 它符合这样的张量计算的传统, 即记向量的分量用上指标, 而记线性型的分量用下指标, 并且对于张量的分量保持类似的约定, 这就使得看一看指标的位置就可以马上了解所考虑的张量的型.

总是 p-重线性的; 当我们用 $u_{h_1}, u_{h_2}, u_{h_{n_h}}$ 表示模 X_h 关于基 $(a_{h_1}, a_{h_2}, \cdots, a_{h_{n_h}})$ 的坐标函数时, 那么显然

$$u_{i_1\cdots i_p}(x_1, \cdots, x_p) = u_{1i_1}(x_1)\cdots u_{pi_p}(x_p),$$

由此得到

$$u_{i_1\cdots i_p} = u_{1i_1} \otimes \cdots \otimes u_{pi_p},$$

这个映射必然是多重线性的 (例 2). 最后留给读者仿照定理 1 和 2 仔细证明 (17) 蕴含 (18).

M 的元素 $c_{i_1\cdots i_p}$ 称为 f 关于 $X_1, \cdots, X_p$ 内所考虑的基的**系数**. 当 $X_1 = \cdots = X_p$ 时, 习惯上在 $X_1, \cdots, X_p$ 内用同样的基 (a_i), 公式 (17) 不变, 而 (18) 写为

$$c_{i_1\cdots i_p} = f(a_{i_1}, \cdots, a_{i_p}), \tag{18'}$$

M 的这些元素称为 f 关于这个基 (a_i) 的系数.

¶例 12　设 T 是有限生成自由模 X 上的一个 $\begin{pmatrix} p \\ q \end{pmatrix}$ 型的张量, 在 X 里选择基 (a_i), 在对偶模 X^* 里选择对偶基 (a^i). 为了对于向量 $x_j \in X$ 和线性型 (或**余向量**) $u_i \in X^*$ 计算 $T(u_1, \cdots, u_p, x_1, \cdots, x_q)$, 令

$$u_h = \sum_{i_h} \alpha_{hi_h} a^{i_h},$$

由此得到 $\alpha_{hi_h} = u_h(a_{i_h})$ 和

$$x_k = \sum_{j_k} \xi_k^{j_k} \cdot a_{j_k};$$

则有

$$T(u_1, \cdots, u_p, x_1, \cdots, x_q) = \sum T_{j_1\cdots j_q}^{i_1\cdots i_p} \alpha_{1i_1} \ldots \alpha_{pi_p} \xi_1^{j_1} \cdots \xi_q^{j_q},$$

其中的常量

$$T_{j_1\cdots j_q}^{i_1\cdots i_p} = T(a^{i_1}, \cdots a^{i_p}, a_{j_1}, \cdots, a_{j_q})$$

称为张量 T 关于 X 的基 (a_i) 的**系数** (或**分量**, 或**坐标**).

假定例如 $p = 2$ 和 $q = 3$, 对于

$$u = \sum \alpha_i a^i, \quad v = \sum \beta_i a^i, \quad x = \sum \xi^i a_i, \quad y = \sum \eta^i a_i, \quad z = \sum \zeta^i a_i,$$

则有

$$T(u, v, x, y, z) = \sum T_{khl}^{ij} \alpha_i \beta_j \xi^k \eta^h \zeta^l,$$

而 T 的分量

$$T_{khl}^{ij} = T(a^i, a^j, a_k, a_h, a_l).$$

我们注意很容易计算张量乘积 (例 7) 的分量. 比如设 U 是一个 $\binom{2}{1}$ 型张量, 而 V 是一个 $\binom{1}{1}$ 型张量, 则 $\mathrm{T} = \mathrm{U} \otimes \mathrm{V}$ 是由

$$\mathrm{T}(u, v, w, x, y) = \mathrm{U}(u, v, x)\mathrm{V}(w, y)$$

给定的 $\binom{3}{2}$ 型张量, T 的分量

$$\mathrm{T}_{hl}^{ijk} = T(a^i, a^j, a^k, a_h, a_l)$$

由关系

$$\mathrm{T}_{hl}^{ijk} = \mathrm{U}_h^{ij}\mathrm{V}_l^k$$

给定. 在一般情形显然有类似的公式.

初学者不要被这些公式和 "纷繁的指标" 所刺激, 本节的结果是些简单的代数等式, 跟公式

$$x(y + z) = xy + xz$$

有同样的深度; 所以所有这些公式本质上都是平凡的, 尽管其外表令人望而生畏, 仅有的困难是选择方便的记号, 而非构思灵巧的推理.

5. 基的变换对于张量分量的影响

设 X 是一个有限生成的自由的 K-模, 考虑 X 的两个基

$$(a_i)_{1 \leqslant i \leqslant n}, \quad (b_\lambda)_{1 \leqslant \lambda \leqslant n};$$

关于第一个基的指标都用拉丁字母, 关于第二个基的指标都用希腊字母. 给定 X 上的一个张量 T, 我们打算通过 T 关于基 (a_i) 的分量计算它关于基 (b_λ) 的分量. 为此令

$$b_\lambda = \sum_i \theta_\lambda^i a_i, \quad a_i = \sum_\lambda \rho_i^\lambda b_\lambda,$$

这样过渡矩阵 (θ_λ^i) 和 (ρ_i^λ) 互为逆矩阵 (§15 第 4 小节的证明). 我们需要从 X^* 的对偶基 (a^i) 到对偶基 (b^λ) 的过渡矩阵. 根据 §16 第 2 小节我们发现对于 $f \in \mathrm{X}^*$ 有

$$f = \sum_i f(a_i) \cdot a^i = \sum_\lambda f(b_\lambda) \cdot b^\lambda.$$

故

$$b^\lambda = \sum_i b^\lambda(a_i) \cdot a^i.$$

而

$$b^\lambda(a_i) = b^\lambda\left(\sum_\mu \rho_i^\mu b_\mu\right) = \sum_\mu \rho_i^\mu b^\lambda(b_\mu),$$

由于根据对偶基的定义

$$b^\lambda(b_\mu)=\begin{cases}0, & \lambda\neq\mu,\\ 1, & \lambda=\mu,\end{cases}$$

于是

$$b^\lambda(a_i)=\rho_i^\lambda.$$

因此得到

$$b^\lambda=\sum_i \rho_i^\lambda a^i,$$

同样有

$$a^i=\sum_\lambda \theta_\lambda^i\cdot b^\lambda.$$

交代了这些, 现在作为例子考虑一个 $\begin{pmatrix}3\\2\end{pmatrix}$ 型的张量. 它关于第一个基的分量是标量

$$\mathrm{T}_{hl}^{ijk}=\mathrm{T}(a^i,a^j,a^k,a_h,a_l),$$

而它关于第二个基的分量是

$$\mathrm{T}_{\alpha\beta}^{\lambda\mu\nu}=\mathrm{T}(b^\lambda,b^\mu,b^\nu,b_\alpha,b_\beta).$$

对于

$$u=\sum\alpha_i a^i,\quad v=\sum\beta_i a^i,\quad w=\sum\gamma_i a^i,\quad x=\sum\xi^i a_i,\quad y=\sum\eta^i a_i,$$

我们曾经得到

$$\mathrm{T}(u,v,w,x,y)=\sum \mathrm{T}_{hl}^{ijk}\alpha_i\beta_j\gamma_k\xi^h\eta^l.$$

用 $b^\lambda,b^\mu,b^\nu,b_\alpha,b_\beta$ 代换 u,v,w,x,y 即得

$$\mathrm{T}_{\alpha\beta}^{\lambda\mu\nu}=\sum_{i,j,k,h,l}\rho_i^\lambda\rho_j^\mu\rho_k^\nu\theta_\alpha^h\theta_\beta^l\mathrm{T}_{hl}^{ijk},\tag{19}$$

这就是要找的关系. 对于其他型的张量显然有类似的公式.



注 3　公式 (19) 经常用作张量的定义, 并且持续到今日. 人们如下进行: 称一个 “几何对象” 为三次共变二次反变的张量, 不明确该对象的 “具体” 性质, 但是约定关于 X 的每个基 (a_i) 它具有 “分量” T_{hl}^{ijk}, 并且根据公式 (19) 这些分量与基相符合. 为了证明这个定义与前文等价, 所有的事情归结为证明如果令 X 的每个基 (a_i) 对应标量

$$\mathrm{T}_{hl}^{ijk},$$

使得对于 X 的任意基 (a_i) 和 (b_λ) 关系 (19) 满足, 则存在 X 上唯一的一个 $\binom{3}{2}$ 型张量 T, 它关于 X 的所有的基 (a_i) 的分量正是和这个基对应的 T^{ijk}_{hl}.

为此, 选定一个基 (a_i), 并且借助这个特殊的基构造张量

$$T(u, v, w, x, y) = \sum T^{ijk}_{hl} \alpha_i \beta_j \gamma_k \xi^h \eta^l,$$

所有的事情归结为证明它关于 X 的所有其他的基 (b_λ) 的分量是关于这个基所对应的 $T^{\lambda\mu\nu}_{\alpha\beta}$. 但是按照前面的计算, T 关于基 (b_λ) 的分量正是标量

$$\sum_{i,j,k,h,l} \rho^\lambda_i \rho^\mu_j \rho^\nu_k \theta^h_\alpha \theta^l_\beta T^{ijk}_{hl},$$

而由于根据假设, 这个表达式正是对应于基 (b_λ) 的分量 $T^{\lambda\mu\nu}_{\alpha\beta}$, 我们的断言被证明.

我们还可以在下列形式下讨论这个推理: 公式 (19) 表示给定了向量 $x, y \in X$ 和余向量 $u, v, w \in X^*$, 表达式

$$\sum T^{ijk}_{hl} \alpha_i \beta_j \gamma_k \xi^h \eta^l$$

借助于 x, y, u, v, w 关于这个基 (a_i) 的坐标来计算, 而事实上与*定义它所用的坐标系无关*. 这恰恰是由于如同 (19) 的公式可以定义一个具有内在意义的对象, 而不依赖构造它所选取的坐标系, 这就使得张量理论在物理学和数学中被引入. 其基本的思想在于坐标系只不过是研究该具有内在意义的对象的工具, 而人们感兴趣的仅仅是这个对象本身.

§21 习题

¶¶ 1. 设 V 是交换域 K 上的 n 维向量空间. 对于任意整数 $p, q \geqslant 0$, 用

$$T^p_q(V)$$

表示由 V 上的 p 次共变 q 次反变张量组成的向量空间.

a) 设 $(a_i)_{1\leqslant i\leqslant n}$ 是 V 的一个基, 而 $(a^i)_{1\leqslant i\leqslant n}$ 是 V^* 的对偶基. 证明元素

$$a_{i_1} \otimes \cdots \otimes a_{i_p} \otimes a^{j_1} \otimes \cdots \otimes a^{j_q} \tag{$*$}$$

(其中的 $i_1, \cdots, j_q$ 取介于 1 和 n 之间的所有值) 组成向量空间 $T^p_q(V)$ 的一个基 [注意: 在 §21 的例 6 会找到把每个元素 $a \in V$ 等同于 $\binom{1}{0}$ 型张量的规则, 为了给上面的表达式 $(*)$ 以意义这个规则是本质的]. 由此推出 $T^p_q(V)$ 的维数是 n^{p+q}.

b) 设 u 是 V 的一个自同构. 给定一个张量 $f \in T^p_q(V)$, 考虑 $(V^*)^p \times V^q$ 上的函数 f', 其定义是, 对于任意 $y_j \in V^*$ 和 $x_i \in V$,

$$f'(y_1, \cdots, y_p, x_1, \cdots, x_q) = f[{}^t u(y_1), \cdots, {}^t u(y_p), u^{-1}(x_1), \cdots, u^{-1}(x_q)].$$

证明 $f' \in \mathrm{T}_q^p(\mathrm{V})$, 并且在 $\mathrm{T}_q^p(\mathrm{V})$ 上这样定义的映射 $f \to f'$ 是线性的. 下面用

$$\mathrm{T}_q^p(u)$$

表示这个映射. 证明对于任何 $u, v \in \mathrm{GL}(\mathrm{V})$ 有

$$\mathrm{T}_q^p(v \circ u) = \mathrm{T}_q^p(v) \circ \mathrm{T}_q^p(v).$$

由此推出映射 $u \to \mathrm{T}_q^p(u)$ 是从 V 的自同构群到 $\mathrm{T}_q^p(\mathrm{V})$ 的自同构群的一个同态. 令

$$u(a_i) = \sum_j \alpha_i^j a_j,$$

则 $(\alpha_i^j)_{1 \leqslant i,j \leqslant n}$ 是 u 关于 V 的基 (a_i) 的矩阵. 计算 $\mathrm{T}_q^p(u)$ 关于 $\mathrm{T}_q^p(\mathrm{V})$ 的问题 a) 所定义的基的矩阵.

c) 设 u 是 V 的一个自同态. 给定一个张量 $f \in \mathrm{T}_q^p(\mathrm{V})$, 定义 $(\mathrm{V}^*)^p \times \mathrm{V}^q$ 上的一个新函数 f'' 如下:

$$\begin{aligned} f''(y_1, \cdots, y_p, x_1, \cdots, x_q) = & \sum_{1 \leqslant i \leqslant p} f[y_1, \cdots, y_{i-1}, {}^t u(y_i), y_{i+1}, \cdots, y_p, x_1, \cdots, x_q] \\ & - \sum_{1 \leqslant j \leqslant q} f[y_1, \cdots, y_p, x_1, \cdots, x_{j-1}, u(x_j), x_{j+1}, \cdots, x_q]. \end{aligned}$$

证明仍然有 $f'' \in \mathrm{T}_q^p(\mathrm{V})$, 并且这样定义的从 $\mathrm{T}_q^p(\mathrm{V})$ 到自身的映射 $f \to f''$ 是线性的. 下面用 $\mathrm{D}_q^p(u)$ 表示这个映射. 证明对于 V 的任何自同态 u, v 有

$$\mathrm{D}_q^p(u+v) = \mathrm{D}_q^p(u) + \mathrm{D}_q^p(v),$$
$$\mathrm{D}_q^p(u \circ v - v \circ u) = \mathrm{D}_q^p(u) \circ \mathrm{D}_q^p(v) - \mathrm{D}_q^p(v) \circ \mathrm{D}_q^p(u).$$

知道了 u 关于 V 的基 (a_i) 的矩阵, 计算 $\mathrm{D}_q^p(u)$ 关于问题 a) 中的基 (*) 的矩阵.

2. 设 V 是交换域 K 上的有限维向量空间, 而 T 是 V 上的一个两次共变和三次反变的张量. 证明存在 V 上唯一的一个一次共变和两次反变张量 U, 它关于 V 的任意的基的分量由关系

$$u_{jk}^i = \sum_{1 \leqslant h \leqslant n} \mathrm{T}_{jkh}^{ih}$$

通过 T 的分量给定. 把 U 和 T 看作多重线性型解释这个运算 (张量 T 关于第二个共变指标和第三个反变指标的**缩并**).

¶3. 设 T 是一个两次共变两次反变的张量. 证明标量

$$\sum_{1 \leqslant i,j \leqslant n} \mathrm{T}_{ji}^{ij}$$

不依赖为了定义它而选择的基.

¶¶4. 设 L 和 M 是交换环 K 上的两个模. 考虑以集合 L × M 作为基础的模

$$\mathrm{N} = \mathrm{K}^{(\mathrm{L} \times \mathrm{M})}$$

(§11, 习题 15). 把 $\mathrm{L}\times\mathrm{M}$ 的每一个元素等同于 N 内的对应元素, 在 N 内考虑由 N 的如下元素生成的子模 N':

$$(\lambda'x'+\lambda''x'',\mu'y'+\mu''y'')-\lambda'\mu'(x',y')-\lambda'\mu''(x',y'')-\lambda''\mu'(x'',y')-\lambda''\mu''(x'',y'').$$

称商模 N/N' 为**模 L 和 M 的张量积**, 记为

$$\mathrm{L}\otimes\mathrm{M}.$$

给定元素 $x\in\mathrm{L}$ 和 $y\in\mathrm{M}$, 用

$$x\otimes y$$

表示 N 的元素 (x,y) 所代表的 $\mathrm{L}\otimes\mathrm{M}=\mathrm{N}/\mathrm{N}'$ 的元素.

a) 证明从 $\mathrm{L}\times\mathrm{M}$ 到 $\mathrm{L}\otimes\mathrm{M}$ 的映射

$$(x,y)\to x\otimes y$$

是双线性的, 并且 "积" $x\otimes y$ 生成模 $\mathrm{L}\otimes\mathrm{M}$.

b) 如果 f 是从 $\mathrm{L}\times\mathrm{M}$ 到任意 K-模 E 内的一个映射. 证明 f 是双线性的, 必须并且只需存在一个线性映射

$$\bar{f}:\mathrm{L}\otimes\mathrm{M}\to\mathrm{E},$$

使得对于任意 $x\in\mathrm{L}$ 和 $y\in\mathrm{M}$ 有

$$f(x,y)=\bar{f}(x\otimes y).$$

映射 $\bar{f}$ 这时是由 f 唯一确定的 (这个结果是模的张量积的基本性质, 用它把对双线性映射的讨论归结为对线性映射的讨论).

c) 假定 L 和 M 是有限生成自由模, 设 $(a_i)_{1\leqslant i\leqslant p}$ 是 L 的一个基, 而 $(b_j)_{1\leqslant j\leqslant q}$ 是 M 的一个基. 证明乘积

$$a_i\otimes b_j\quad(1\leqslant i\leqslant p,1\leqslant j\leqslant q)$$

是 $\mathrm{L}\otimes\mathrm{M}$ 的一个基. 由此推出, 如果 K 是一个域, 则有

$$\dim(\mathrm{L}\otimes\mathrm{M})=\dim(\mathrm{L})\cdot\dim(\mathrm{M}).$$

d) 证明存在唯一的一个从 $\mathrm{L}\otimes\mathrm{M}$ 到 $\mathrm{M}\otimes\mathrm{L}$ 上的同构, 使得对于任意 $x\in\mathrm{L}$ 和 $y\in\mathrm{M}$, $x\otimes y$ 映射到 $y\otimes x$.

e) 设 $\mathrm{L},\mathrm{M},\mathrm{L}',\mathrm{M}'$ 是 K 上的四个模, 考虑同态

$$u:\mathrm{L}\to\mathrm{L}'\quad 和\quad v:\mathrm{M}\to\mathrm{M}'.$$

证明存在唯一的一个同态

$$f:\mathrm{L}\otimes\mathrm{M}\to\mathrm{L}'\otimes\mathrm{M}',$$

使得对于任意 $x\in\mathrm{L}$ 和 $y\in\mathrm{M}$ 有

$$f(x\otimes y)=u(x)\otimes v(y)$$

(注意右端是 x 和 y 的双线性函数). 我们说 f 是**同态 u 和 v 的张量积**, 并且一般记作 $u\otimes v$. [这个传统的记号可能会引起混淆, 因为它还用来表示 K-模

$$\mathrm{Hom(L, L')\otimes Hom(M, M')}$$

的一个元素. 在实际中, 几乎从不考虑后一个张量积, $u\otimes v$ 总是具有前面的含义.]

f) 假定 $\mathrm{L, M, L', M'}$ 是有限生成自由模. 选择这些模的基, 从而 (根据前面的问题 c)) 也就选择了 $\mathrm{L\otimes M}$ 和 $\mathrm{L'\otimes M'}$ 的基, 用 u 和 v 关于在 $\mathrm{L, M, L', M'}$ 中选择的基的矩阵计算 $u\otimes v$ 关于 $\mathrm{L\otimes M}$ 和 $\mathrm{L'\otimes M'}$ 中的相应的基的矩阵. [所得到的结果引导出**两个矩阵的张量积**的概念. 设

$$\mathrm{A}=(a_{ij})_{1\leqslant i\leqslant p, 1\leqslant j\leqslant q}\quad 和\quad \mathrm{B}=(b_{kl})_{1\leqslant k\leqslant m, 1\leqslant l\leqslant n}$$

是其元素在一个交换环 K 内的矩阵. 称如下定义的 pm 行 qn 列矩阵为 A 和 B 的张量积: 借助序偶 $(i,k)(1\leqslant i\leqslant p, 1\leqslant k\leqslant m)$ 给行编号, 借助序偶 $(j,l)(1\leqslant j\leqslant q, 1\leqslant l\leqslant n)$ 给列编号; 这一点约定好了, 张量积

$$\mathrm{A\otimes B}$$

位于指标为 (i,k) 的行和指标为 (j,l) 的列的交叉处的元素是

$$a_{ij}b_{kl}.$$

例如

$$\begin{pmatrix} a & b\\ c & d\end{pmatrix}\otimes\begin{pmatrix} u & v\\ u' & v'\\ u'' & v''\end{pmatrix}=\begin{pmatrix} au & av & bu & bv\\ cu & cv & du & dv\\ au' & av' & bu' & bv'\\ cu' & cv' & du' & dv'\\ au'' & av'' & bu'' & bv''\\ cu'' & cv'' & du'' & dv''\end{pmatrix}.$$

为了对序偶 (i,k) 排序, 我们采用字典排序法, 约定如果 $i<k$, 或 $i=k$ 且 $j<h$, 则 (i,j) 先于 (k,h).]

g) 设 $\mathrm{L, L', L'', M, M'}$ 和 $\mathrm{M''}$ 是环 K 上的六个模, 取同态

$$u':\mathrm{L}\to\mathrm{L'},\quad u'':\mathrm{L'}\to\mathrm{L''},\quad v':\mathrm{M}\to\mathrm{M'},\quad v'':\mathrm{M'}\to\mathrm{M''}.$$

证明我们有

$$(u''\circ u')\otimes(v''\circ v')=(u''\otimes v'')\circ(u'\otimes v').$$

由此推出, 如果 $\mathrm{A', A'', B', B''}$ 是其元素在 K 内的矩阵, 公式

$$(\mathrm{A''A'})\otimes(\mathrm{B''B'})=(\mathrm{A''\otimes B''})\cdot(\mathrm{A'\otimes B'})$$

成立, 只要它有意义.

h) 设 L, M 和 N 是三个模. 证明存在唯一的一个同构

$$(\mathrm{L\otimes M})\otimes\mathrm{N}\to\mathrm{L}\otimes(\mathrm{M\otimes N}),$$

对于任意 $x \in \mathrm{L}, y \in \mathrm{M}$ 和 $z \in \mathrm{N}$, 它映射 $(x \otimes y) \otimes z$ 到 $x \otimes (y \otimes z)$. [张量积的 "结合律". 在实际中, 对于 $(x \otimes y) \otimes z$ 和 $x \otimes (y \otimes z)$ 不做任何区分, 都写成 $x \otimes y \otimes z$.]

i) 设 $\mathrm{M}_1, \cdots, \mathrm{M}_p$ 是 K 上的模, 组成模

$$\mathrm{M}_1 \otimes \mathrm{M}_2 \otimes \cdots \otimes \mathrm{M}_p = \mathrm{M}_1 \otimes (\mathrm{M}_2 \otimes \cdots \otimes \mathrm{M}_p)$$

(关于 p 归纳定义). 设 f 是从 $\mathrm{M}_1 \times \cdots \times \mathrm{M}_p$ 到一个 K-模 $\mathbf{N}$ 的映射. 证明 f 是 p 重线性的, 必须并且只需存在模的一个同态

$$\bar{f}: \mathrm{M}_1 \otimes \mathrm{M}_2 \otimes \cdots \otimes \mathrm{M}_p \to \mathrm{N},$$

使得对于任意 $x_i \in \mathrm{M}_i$ 有

$$f(x_1, \cdots, x_n) = \bar{f}(x_1 \otimes \cdots \otimes x_n).$$

线性映射 $\bar{f}$ 则完全由 f 确定 (关于 p 归纳推理, 赋予出现在 f 中的变量之一一个固定值).

j) 设 M 是一个 K-模, 考虑模

$$\mathrm{M} \otimes \mathrm{M} \otimes \mathrm{M}^*,$$

其中 M^* 是 M 的对偶模. 借助前一个问题, 证明存在唯一的一个从 $\mathrm{M} \otimes \mathrm{M} \otimes \mathrm{M}^*$ 到二次共变一次反变的张量的模内的同态

$$j : \mathrm{M} \otimes \mathrm{M} \otimes \mathrm{M}^* \to \mathrm{T}_1^2(\mathrm{M}),$$

它映射元素

$$x \otimes y \otimes u \in \mathrm{M} \otimes \mathrm{M} \otimes \mathrm{M}^*$$

到元素

$$x \otimes y \otimes u \in \mathrm{T}_1^2(\mathrm{M}).$$

[我们回忆在 §21 的例 7 里提到, 最后一个表达式的意思是 $\mathrm{M}^* \otimes \mathrm{M}^* \otimes \mathrm{M}$ 上的一个三重线性型 $x \otimes y \otimes u$, 它在 $(f, g, z) \in \mathrm{M}^* \times \mathrm{M}^* \times \mathrm{M}$ 的值是 K 的元素 $f(x)g(y)u(z)$.]

证明如果 M 是有限生成自由模, 则 j 是双射. 推广这个结果, 其中用 $\mathrm{M} \otimes \cdots \otimes \mathrm{M} \otimes \mathrm{M}^* \otimes \cdots \otimes \mathrm{M}^*$ (p 个因子 M 和 q 个因子 M^*) 代替 $\mathrm{M} \otimes \mathrm{M} \otimes \mathrm{M}^*$, 并且用前面习题 1 已经给了定义的 $\mathrm{T}_q^p(\mathrm{M})$ 代替 $\mathrm{T}_1^2(\mathrm{M})$.

k) 取 $\mathrm{K} = \mathbf{Z}$ 和 $\mathrm{M} = \mathbf{Z}/p\mathbf{Z}$, 证明 $\mathrm{T}_0^2(\mathrm{M})$ 缩减为 0 模, 但是 $\mathrm{M} \otimes \mathrm{M}$ 则不然 [考虑从 $\mathrm{M} \times \mathrm{M}$ 到 M 内的映射 $(x, y) \to xy$, 这是模 p 整数的乘法]. 由此推出, 在这种情形前一个问题的映射 j 不是双射 (并且甚至是零).

[在这个习题中定义的两个模的张量积的概念比在 §21 定义的有用得多, 除了问题涉及的是有限生成的自由模. 其理由是, 在 §21 中的张量用于研究到基础环 K 内的多重线性映射, 而在习题 4 中的张量适合于研究到任意 K-模内的多重线性映射. 而从本题的 k) 看到, 有可能前一个张量积恒等于零, 而后一个张量积则不然. 最后注意, 矩阵或有限维向量空间的张量积本质上起始于 Kronecker, 基于这个理由, 一些著作者称它为 **Kronecker 积**.]

§22 交错双线性和三重线性映射

1. 交错双线性映射

设 X 和 M 是交换环 K 上的模. 称一个双线性映射

$$f : \mathrm{X} \times \mathrm{X} \to \mathrm{M}$$

是**交错的**, 如果有

$$f(x,x) = 0 \quad \text{对于所有 } x \in \mathrm{X}. \tag{1}$$

于是, 对于任意 $x, y \in \mathrm{X}$ 有

$$0 = f(x+y, x+y) = f(x,x) + f(x,y) + f(y+x) + f(y,y) = f(x,y) + f(y+x),$$

因此交错双线性映射满足

$$f(y,x) = -f(x,y), \quad \text{任意 } x, y \in \mathrm{X}. \tag{2}$$



注 1 显然反之 (2) 蕴含

$$2 \cdot f(x,x) = 0 \quad \text{对于所有 } x \in \mathrm{X}.$$

如果对于 $m \in \mathrm{M}$ 关系 $2m = 0$ 蕴含 $m = 0$ (如果域 K 是 "特征" 异于 2 的域: 参见 §30, 第 6 小节), 那么关系 (2) 就刻画了交错双线性映射的特征. 反之, 如果 $\mathrm{M} = \mathrm{K}$ 是模 2 整数环, 那么这时关系 (2) 写作 $f(y,x) = f(x,y)$, 不足以蕴含 f 是交错的. 这种状况在初等实践中是很少遇到的.

显然所有交错双线性映射的线性组合还是交错双线性映射, 即交错双线性映射组成从 $\mathrm{X} \times \mathrm{X}$ 到 M 内的所有双线性映射构成的模 $\mathscr{L}(\mathrm{X}, \mathrm{X}; \mathrm{M})$ 的一个子模.

例 1 取 K =**R**, 取通常的三维空间为 X, 那么从 $\mathrm{X} \times \mathrm{X}$ 到 X 内由

$$f(x,y) = x \wedge y$$

定义的映射 (向量积; 参见 §21, 例 4) 是交错双线性的.

例 2 设 f 是从 $\mathrm{X} \times \mathrm{X}$ 到 M 内的双线性映射, 则由

$$g(x,y) = f(x,y) - f(y,x)$$

定义的映射是交错双线性的. 一个特殊情形是, 如果 u 和 v 是 X 上的线性型, 那么从 $\mathrm{X} \times \mathrm{X}$ 到 M 内的映射

$$g(x,y) = u(x)v(y) - u(y)v(x)$$

是 X × X 上的一个交错双线性型; 称为**线性型** u **和** v **的外积**, 并且记作

$$u \wedge v.$$

例如取 K =**R**, 而 X 是前一个例子中的模, 并且

$$u(x) = (a|x), \quad v(x) = (b|x),$$

其中 a 和 b 是固定的向量. 对于 $g = u \wedge v$ 我们有公式

$$g(x, y) = (a \wedge b|x|y),$$

这就是说有等式

$$(a|x)(b|y) - (a|y)(b|x) = (a \wedge b|x|y) = (a|b|x \wedge y).$$

留给读者仔细证明这个结果, 或者采用直接的几何推理, 或者通过直角坐标系进行计算 (甚至可以在任意坐标系里计算).

2. 有限生成自由模的情形

对于交错双线性映射有类似于 §21 的定理 1 的结果:

定理 1 设 X 和 M 是交换环 K 上的模, 假定 X 是有限生成自由的. 设 $(a_i)_{1 \leqslant i \leqslant n}$ 是 X 的一个基. 从 X × X 到 M 内的双线性映射是交错的, 必须并且只需它的关于基 (a_i) 的系数

$$c_{ij} = f(a_i, a_j)$$

满足关系

$$c_{ii} = 0, \quad c_{ij} + c_{ji} = 0. \tag{3}$$

当条件 (3) 满足时, 对于任意

$$x = \sum_i \xi_i a_i, \quad y = \sum_i \eta_i a_i$$

有

$$f(x, y) = \sum_{i<j} c_{ij}(\xi_i \eta_j - \xi_j \eta_i). \tag{4}$$

对于 $x = a_i$ 写出 (1), 对于 $x = a_i$ 和 $y = a_j$ 写出 (2), 显然得到关系 (3). 反之假定 (3) 成立, 在从 §21 的定理 1 得到的公式

$$f(x, y) = \sum_{1 \leqslant i, j \leqslant n} \xi_i \eta_j c_{ij}$$

里, 满足 $i=j$ 的项是零, 根据 i 和 j 的大小把其余的项分组得到

$$f(x,y)=\sum_{i<j}\xi_i\eta_j c_{ij}+\sum_{i>j}\xi_i\eta_j c_{ij}.$$

在第二个和式里, 字母 i 和 j 换成 j 和 i (仅涉及记号变换), 根据 (3) 我们得到

$$f(x,y)=\sum_{i<j}\xi_i\eta_j c_{ij}+\sum_{i<j}\xi_j\eta_i c_{ji}=\sum_{i<j}\xi_i\eta_j c_{ij}-\sum_{i<j}\xi_j\eta_i c_{ij},$$

由此即得关系 (4), 而这个关系表明 f 是交错的, 因为当 $x=y$ 时, 差 $\xi_i\eta_j-\xi_j\eta_i$ 显然是零. 这就完成了证明.



注 2 用 u_i 表示 X 关于基 (a_i) 的坐标函数, 根据例 2 我们有

$$\xi_i\eta_j-\xi_j\eta_i=u_i(x)u_j(y)-u_j(x)u_i(y)=(u_i\wedge u_j)(x,y),$$

因此 (4) 可以写成

$$f(x,y)=\sum_{i<j}u_{ij}(x,y)c_{ij},$$

其中 $u_{ij}=u_i\wedge u_j$, 并且 f 的这个分解是唯一的, 因为它蕴含 $c_{ij}=f(a_i,a_j)$.

在 $\mathrm{M}=\mathrm{K}$ 这个特殊情形, 前面的公式写成

$$f=\sum_{i<j}\gamma_{ij}u_{ij},\quad \text{其中 } \gamma_{ij}=f(a_i,a_j),$$

因此推出 $n(n-1)/2$ 个交错双线性型 $u_{ij}=u_i\wedge u_j (1\leqslant i<j\leqslant n)$ 组成 $\mathbf{X}\times\mathbf{X}$ 上的所有交错双线性型的模的一个基.

利用已经在 §15 第 3 小节引进的记号

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}=ad-bc,$$

则有

$$f(x,y)=\sum_{i<j}\gamma_{ij}\begin{vmatrix} \xi_i & \eta_i \\ \xi_j & \eta_j \end{vmatrix},\quad \text{其中 } \gamma_{ij}=f(a_i,a_j). \tag{5}$$

当 $n=1$ 时, 显然 f 恒等于零, 这是因为

$$f(x,y)=f(\xi_1a_1,\eta_1a_1)=\xi_1\eta_1f(a_1,a_1)=0.$$

当 $n=2$ 时, 公式 (5) 缩减为

$$f(x,y)=\gamma_{12}\begin{vmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \end{vmatrix},\quad \text{其中 } \gamma_{12}=f(a_1,a_2). \tag{6}$$

函数

$$D(x, y) = \begin{vmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \end{vmatrix} = \xi_1\eta_2 - \xi_2\eta_1$$

称为**向量 x 和 y 关于 X 的基** (a_1, a_2) **的行列式**. 这是 X 上的一个交错双线性型, 使得

$$D(a_1, a_2) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1,$$

这个性质刻画了 D 的特征, 这是因为 (6) 可以写成

$$f = f(a_1, a_2) \cdot D,$$

如果 $D(a_1, a_2) = 1$, 则有 $f = D$.

当 $n = 3$ 时, 公式 (5) 写成

$$f(x, y) = \gamma_{23} \begin{vmatrix} \xi_2 & \eta_2 \\ \xi_3 & \eta_3 \end{vmatrix} + \gamma_{31} \begin{vmatrix} \xi_3 & \eta_3 \\ \xi_1 & \eta_1 \end{vmatrix} + \gamma_{12} \begin{vmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \end{vmatrix}, \tag{7}$$

在这种情形, $X \times X$ 上的交错双线性型依赖于三个任意常量

$$\gamma_{23} = f(a_2, a_3), \quad \gamma_{31} = f(a_3, a_1), \quad \gamma_{12} = f(a_1, a_2).$$

注 3 当 K =**R** 并且 X 是通常三维空间时, 假定 X 的基 $(a_i)_{1 \leqslant i \leqslant 3}$ 是正交规范的. 考虑向量



$$u = \gamma_{23}a_1 + \gamma_{31}a_2 + \gamma_{12}a_3, \tag{8}$$

关系 (7) 表明 $f(x, y)$ 是 u 和向量 $x \wedge y$ 的标量积, 即 (§21, 例 10)

$$f(x, y) = (u|x|y).$$

反之, 显然所有由这种类型的公式给定的函数是 X 上的一个交错双线性型.

于是可以把 f 等同于向量 u, 这就解释了为什么交错双线性型一般不介入到初等几何或物理学中. 无论如何必须注意到 f 仅在直角坐标系的情形 (而直角坐标系要假定选择了一个长度单位) 等同于向量 (8); 更精确地说, 向量 (8) 在从一个直角坐标系过渡到一个直角坐标系的范围内不依赖基 (a_i); 但是如果我们准许 (a_i) 在 X 的所有的基的集合内变换, 那么向量 (8) 不仅仅依赖于 f, 还依赖所考虑的基 (a_i). 这来源于事实: 坐标变换公式对于一个 $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ 型的张量跟对于一个向量即一个 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 型张量是不同的.

3. 交错三重线性映射

给定在一个交换环上的模 X 和 M, 称一个三重线性映射

$$f: \mathrm{X} \times \mathrm{X} \times \mathrm{X} \to \mathrm{M}$$

是交错的, 如果向量 x, y, z 中一旦有两个相等, 表达式 $f(x, y, z)$ 就等于零, 即对于任意 $x, y, z \in \mathrm{X}$, 有

$$f(x, y, y) = f(x, y, x) = f(x, x, z) = 0. \tag{9}$$

因此对于所有 $a \in \mathrm{X}$, 双线性函数 $f(a, y, z), f(x, a, z)$ 和 $f(x, y, a)$ 是交错的, 从而有关系

$$f(x, y, z) = -f(x, z, y), \quad f(x, y, z) = -f(z, y, x), \quad f(x, y, z) = -f(y, x, z),$$

这些显然可以归结为关系

$$f(x, y, z) = f(y, z, x) = f(z, x, y) = -f(x, z, y) = -f(y, x, z) = -f(z, y, x). \tag{10}$$

如果对于一个 $m \in \mathrm{M}$, $2m = 0$ 蕴含 $m = 0$, 那么关系 (10) 蕴含关系 (9), 故可以刻画交错三重线性函数的特征.

例 3　如果取 $\mathrm{K} = \mathbf{R}$, 取通常的三维空间作为 X, 混合积 $(x|y|z)$ 是 X 上的交错三重线性型.

例 4　X 和 K 是任意的, 取一个三重线性映射 g, 并且令

$$\begin{aligned} f(x, y, z) = &\, g(x, y, z) + g(y, z, x) + g(z, x, y) \\ &- g(x, z, y) - g(y, x, z) - g(z, y, x); \end{aligned}$$

那么立刻看出 f 是一个交错三重线性映射.

考虑一个特殊情形, 设 u, v, w 是 X 上的线性型, 那么

$$\begin{aligned} f(x, y, z) = &\, u(x)v(y)w(z) + u(y)v(z)w(x) + u(z)v(y)w(y) \\ &- u(x)v(z)w(y) - u(u)v(x)w(z) - u(z)v(y)w(x) \end{aligned}$$

是 $\mathrm{X} \times \mathrm{X} \times \mathrm{X}$ 上的一个交错三重线性型, 称为线性型 u, v, w 的**外积**, 并且用记号

$$u \wedge v \wedge w$$

表示. 如果 u, v, w 中有两个相等, 则 $u \wedge v \wedge w = 0$, 并且有

$$u \wedge v \wedge w = v \wedge w \wedge u = w \wedge u \wedge v = -u \wedge w \wedge v = -v \wedge u \wedge w = -w \wedge v \wedge u.$$

例 5 设 u 是 X 上的一个线性型, 而 f 是模 X 上的一个交错双线性型, 那么函数

$$g(x,y,z)=u(x)f(y,z)+u(y)f(z,x)+u(z)f(x,y)$$

是 $\mathrm{X}\times\mathrm{X}\times\mathrm{X}$ 上的一个交错三重线性型, 称为**线性型 u 乘以交错双线性型 f 的外积**, 并且用记号

$$u\wedge f$$

表示. 如果 u,v,w 是 X 上的线性型, 则有

$$(u\wedge v)\wedge w=u\wedge(v\wedge w).$$

同样, 用

$$f\wedge u$$

表示交错三重线性型

$$f(x,y)u(z)+f(y,z)u(x)+f(z,x)u(y),$$

我们有

$$u\wedge f=f\wedge u$$

(而当 u 和 v 是两个线性型时, $u\wedge v=-v\wedge u$).

4. 关于一个基的展开

对于交错三重线性映射有类似于定理 1 的定理:

定理 2 设 X 和 M 是一个交换环 K 上的模. 假定 X 是有限生成自由的, 并且 $(a_i)_{1\leqslant i\leqslant n}$ 是 X 的一个基. 从 $\mathrm{X}\times\mathrm{X}\times\mathrm{X}$ 到 M 内的三重线性映射是交错的, 必须并且只需它关于所考虑的基 (a_i) 的系数

$$c_{ijk}=f(a_i,a_j,a_k)$$

满足关系

$$c_{ijj}=c_{iji}=c_{iik}=0, \tag{11}$$

$$c_{ijk}=c_{jki}=c_{kij}=-c_{ikj}=-c_{jik}=-c_{kji}; \tag{12}$$

这些条件满足时对于任意向量

$$x=\sum\xi_i a_i,\quad y=\sum\eta_i a_i,\quad z=\sum\zeta_i a_i$$

有

$$f(x,y,z)=\sum_{i<j<k}c_{ijk}\begin{vmatrix}\xi_i & \eta_i & \zeta_i\\ \xi_j & \eta_j & \zeta_j\\ \xi_k & \eta_k & \zeta_k\end{vmatrix}. \tag{13}$$

在 (13) 中我们采用了类似于二阶行列式并且已经在 §21 例 10 引进的记号

$$\begin{vmatrix} \xi_i & \eta_i & \zeta_i \\ \xi_j & \eta_j & \zeta_j \\ \xi_k & \eta_k & \zeta_k \end{vmatrix} = \xi_i\eta_j\zeta_k + \xi_j\eta_k\zeta_i + \xi_k\eta_i\zeta_j - \xi_i\eta_k\zeta_j - \xi_j\eta_i\zeta_k - \xi_k\eta_j\zeta_i. \tag{14}$$

为了证明定理 2, 应当首先指出对于所有交错三重线性映射有 (11) 和 (12). 这只要对于 $x = a_i, y = a_j, z = a_k$ 写出关系 (9) 和 (10) 就可以了.

反之, 假定 (11) 和 (12) 满足. 在由 §21, 定理 2 推出的公式

$$f(x,y,z) = \sum_{1\leqslant i,j,k\leqslant n} c_{ijk}\xi_i\eta_j\zeta_k$$

中, 仅有的可能的非零项是指标 i,j,k 两两不同的那些项. 把三元组 (i,j,k) 按照 i,j,k 相对大小分类, 就得到六个部分和的分解, 即

$$\begin{aligned} f(x,y,z) = &\sum_{i<j<k} c_{ijk}\xi_i\eta_j\zeta_k + \sum_{j<k<i} c_{ijk}\xi_i\eta_j\zeta_k + \sum_{k<i<j} c_{ijk}\xi_i\eta_j\zeta_k \\ &+ \sum_{i<k<j} c_{ijk}\xi_i\eta_j\zeta_k + \sum_{j<i<k} c_{ijk}\xi_i\eta_j\zeta_k + \sum_{k<j<i} c_{ijk}\xi_i\eta_j\zeta_k. \end{aligned}$$

对于每个和式, 可以改变指标的记号, 使得每个和式都是对于满足 $i<j<k$ 的三元组 (i,j,k) 求和 (例如, 在第二个和式里, j 换成 i, k 换成 j, i 换成 k). 这样做就会发现

$$\begin{aligned} f(x,y,z) = &\sum_{i<j<k} c_{ijk}\xi_i\eta_j\zeta_k + \sum_{i<j<k} c_{kij}\xi_k\eta_i\zeta_j + \sum_{i<j<k} c_{jki}\xi_j\eta_k\zeta_i \\ &+ \sum_{i<j<k} c_{ikj}\xi_i\eta_k\zeta_j + \sum_{i<j<k} c_{jik}\xi_j\eta_i\zeta_k + \sum_{i<j<k} c_{kji}\xi_k\eta_j\zeta_i; \end{aligned}$$

考虑到关系 (12), 把指标同为 i,j,k 的项合并在一起便得到

$$f(x,y,z) = \sum_{i<j<k} c_{ijk}(\xi_i\eta_j\zeta_k + \xi_k\eta_i\zeta_j + \xi_j\eta_k\zeta_i - \xi_i\eta_k\zeta_j - \xi_j\eta_i\zeta_k - \xi_k\eta_j\zeta_i),$$

这恰好是等式 (13).

还要验证 f 事实上是交错的, 为此只需验证三重线性型

$$u_{ijk}(x,y,z) = \xi_i\eta_j\zeta_k + \xi_k\eta_i\zeta_j + \xi_j\eta_k\zeta_i - \xi_i\eta_k\zeta_j - \xi_j\eta_i\zeta_k - \xi_k\eta_j\zeta_i$$

是交错的. 而如果用 u_i 表示模 X 关于基 (a_i) 的第 i 个坐标函数, 那么就有

$$\begin{aligned} u_{ijk}(x,y,z) = &u_i(x)u_j(y)u_k(z) + u_i(y)u_j(z)u_k(x) + u_i(z)u_j(x)u_k(y) \\ &-u_i(x)u_j(z)u_k(y) - u_i(y)u_j(x)u_k(z) - u_i(z)u_j(y)u_k(x), \end{aligned}$$

即

$$u_{ijk} = u_i \wedge u_j \wedge u_k,$$

这就证明了 (例 5) 所提及的表达式是交错的. 因此定理 2 证明完毕.

给定一个系数在一个交换环 K 内的三阶矩阵

$$\mathrm{A} = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix},$$

称元素

$$ab'c'' + bc'a'' + ca'b'' - ac'b'' - ba'c'' - cb'a''$$

为 A 的**行列式**, 或者用定理 2 的叙述中用到的记号

$$\begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix}$$

表示它, 或者用记号

$$\det(\mathrm{A})$$

表示它. 注意有公式

$$\begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix} = a \cdot \begin{vmatrix} b' & b'' \\ c' & c'' \end{vmatrix} - a' \cdot \begin{vmatrix} b & b'' \\ c & c'' \end{vmatrix} + a'' \cdot \begin{vmatrix} b & b' \\ c & c' \end{vmatrix}.$$

例 6 当 $n \leqslant 2$, 指标 i, j, k 不可能两两不同, 系数 c_{ijk} 全是零, 随后仅有的交错三重线性映射是 $f = 0$.

例 7 假定 $n = 3$, 那么 (13) 缩减为

$$f(x, y, z) = c_{123} \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix}, \quad 其中\ c_{123} = f(a_1, a_2, a_3).$$

表达式

$$\mathrm{D}(x, y, z) = \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix}$$

称为**向量** x, y, z **关于基** (a_1, a_2, a_3) **的行列式**. 事实上它的值依赖模 X 的基的选取. 函数 $\mathrm{D}(x, y, z)$ 是 X 上的一个交错三重线性型, 对于它有

$$\mathrm{D}(a_1, a_2, a_3) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1,$$

这个关系在 X 上的交错三重线性型中刻画了 D 的特征. 其实, 对于一个这样的线性型, 根据定理 2 有

$$f = f(a_1, a_2, a_3) \cdot \mathrm{D},$$

于是如果 f 在基向量上等于 1, 则 $f = \mathrm{D}$.

例 8 取 $\mathrm{K} = \mathbf{R}$, 取三维通常空间为 X, 考虑混合积 $(x|y|z)$. 这是 X 上的一个交错三重线性型, 并且关于 X 的任何基有

$$(x|y|z) = (a_1|a_2|a_3) \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix},$$

我们重新得到 §21 例 10 的结果.

交错三重线性型和三阶行列式的研究除前面提到的以外还包括许多其他的结果, 但是我们不在这里叙述了. 在下面两节我们要把它们推广到交错 p-线性型和任意阶的行列式.

§22 习题

计算下列行列式:

1. $\begin{vmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{vmatrix}$.

2. $\begin{vmatrix} 2\sin t\cos t & 2\sin^2 t - 1 \\ 2\cos^2 t - 1 & 2\sin t\cos t \end{vmatrix}$.

3. $\begin{vmatrix} 4 & -3 & 5 \\ 3 & -2 & 8 \\ 1 & -7 & -5 \end{vmatrix}$.

4. $\begin{vmatrix} 3 & 4 & -5 \\ 8 & 7 & -2 \\ 2 & -1 & 8 \end{vmatrix}$.

5. $\begin{vmatrix} x^2+1 & xy & xz \\ xy & y^2+1 & yz \\ xz & yz & z^2+1 \end{vmatrix}$.

6. $\begin{vmatrix} \cos a & \sin a\cos b & \sin a\sin b \\ -\sin a & \cos a\cos b & \cos a\sin b \\ 0 & -\sin b & \cos b \end{vmatrix}$.

7. $\begin{vmatrix} 1 & 0 & 1+i \\ 0 & 1 & i \\ 1-i & -i & 1 \end{vmatrix}$.

8. $\begin{vmatrix} 1 & 1 & 1 \\ 1 & z & z^2 \\ 1 & z^2 & z \end{vmatrix}$, 其中 $z = \cos(4\pi/3) + i\sin(4\pi/3)$.

9. $\begin{vmatrix} \sin^2 a & \cos 2a & \cos^2 a \\ \sin^2 b & \cos 2b & \cos^2 b \\ \sin^2 c & \cos 2c & \cos^2 c \end{vmatrix}$.

10. $\begin{vmatrix} 1 & a & a^4 \\ 1 & b & b^4 \\ 1 & c & c^4 \end{vmatrix}$.

11. $\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix}$.

12. $\begin{vmatrix} x+a & b & c \\ a & x+b & c \\ a & b & x+c \end{vmatrix}$.

13. $\begin{vmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{vmatrix}$.

14. $\begin{vmatrix} a & b & c \\ a & x & c \\ a & b & x \end{vmatrix}$.

15. $\begin{vmatrix} ab & ab' & ab'' \\ a'b & a'b' & a'b'' \\ a''b & a''b' & a''b'' \end{vmatrix}$.

16. $\begin{vmatrix} 1 & x & x^2 \\ a & 1 & x \\ b & c & 1 \end{vmatrix}$.

¶17. 设 V 是交换域 K 上的 n 维向量空间, 而 f 是 V 上的一个交错双线性型. 假定仅有的使得

$$f(a,x)=0 \quad 对于所有\ x \in \mathrm{X}$$

是 $a=0$ (这时称 f 是**非退化的**).

a) 设 A 是由 f 关于 V 的基 (a_i) 的系数 $\alpha_{ij}=f(a_i,a_j)$ 组成的矩阵. 证明 f 是非退化的, 必须并且只需 A 是可逆的 (利用 §20 的定理 2).

b) 设 $a,b\in \mathrm{V}$ 使得 $f(a,b)\neq 0$. 令 $f_a(x)=f(a,x)$ 和 $f_b(x)=f(b,x)$, 证明 f_a 和 f_b 是 V 上的不成比例的线性型, 并且使得

$$f(a,x)=f(b,x)=0 \tag{$*$}$$

的 $x\in \mathrm{V}$ 组成 V 的 $n-2$ 维子空间.

c) 保留 b) 的假设, 证明 V 是由 a 和 b 所生成的平面和 $(*)$ 的解的子空间 V' 的直和. 证明 f 在 V' 的限制是非退化的.

d) 关于 n 进行归纳推理, 证明 i) 如果在 V 上存在一个非退化的交错双线性型, 则 V 的维数是偶数. ii) 如果 f 是 $2p$ 维向量空间 V 上的一个非退化的交错双线性型, 则存在 V 的一个基, 使得 f 关于这个基的矩阵是

$$\begin{pmatrix} 0_p & 1_p \\ -1_p & 0_p \end{pmatrix},$$

其中的 0_p 是 p 行 p 列的零矩阵.

e) 一个其元素在 K 内的方阵 $\mathrm{A}=(\alpha_{ij})_{1\leqslant i,j\leqslant n}$ 称为**交错的**或**反对称的**, 如果它满足关系

$$\alpha_{ii}=0, \quad \alpha_{ij}+\alpha_{ji}=0.$$

证明如果 n 是奇数, 这样的矩阵不可能是可逆的. 如果 A 是可逆的, 并且 $n=2p$, 则存在一个矩阵 $\mathrm{U}\in \mathrm{GL}(n,\mathrm{K})$, 使得

$$\mathrm{U}\mathrm{A}{}^t\mathrm{U}=\begin{pmatrix} 0_p & 1_p \\ -1_p & 0_p \end{pmatrix}.$$

f) 证明对于任意 $x, y, z \in \mathrm{K}$ 有

$$\begin{vmatrix} 0 & x & z \\ -x & 0 & y \\ -z & -y & 0 \end{vmatrix} = 0.$$

18. 设 f, g, h 是交换域 K 上的有限维向量空间 V 上的三个线性型. 证明 f, g, h 是线性无关的, 必须并且只需

$$f \wedge g \wedge h \neq 0.$$

19. 设 V 是交换域 K 上的 n 维向量空间, $(a_i)_{1 \leqslant i \leqslant n}$ 是 V 的一个基, 而 f, g, h 是 V 上的三个线性型. 证明 $f \wedge g \wedge h$ 关于基 (a_i) 的系数是标量

$$\alpha_{ijk} = \begin{vmatrix} f(a_i) & f(a_j) & f(a_k) \\ g(a_i) & g(a_j) & g(a_k) \\ h(a_i) & h(a_j) & h(a_k) \end{vmatrix}.$$

20. 设 V 是交换域 K 上的有限维向量空间, (a_i) 是 V 的一个基, f 是 V 上的交错双线性型, g 是 V 上的线性型. 证明交错三重线性型 $f \wedge g$ 关于基 (a_i) 的系数是标量

$$\alpha_{ijk} = f(a_i, a_j)g(a_k) + f(a_j, a_k)g(a_i) + f(a_k, a_i)g(a_j).$$

21. 设 V 是交换域上的三维向量空间, 而 x, y, z 是 V 的三个元素.

a) 如果 x, y, z 是线性相关的, 则对于 V 上的所有交错三重线性型有 $f(x, y, z) = 0$ (向量中的一个用其余的向量表示).

b) 如果 x, y, z 是线性无关的, 则对于 V 上的所有非零交错三重线性型 f 有 $f(x, y, z) \neq 0$ (注意 x, y, z 组成 V 的基).

c) 设 a, b, c 是 V 的一个基. x, y, z 是线性无关的, 必须并且只需它们关于基 a, b, c 的坐标的行列式不是零 (利用问题 a) 和 b) 以及 §22 的例 7).

(这个习题的结果将在下一节里推广, 但是还是建议读者首先考察三维空间的情形, 因为在实际中这是极其重要的情形.)

22. 向量

$$(2, -3, 1), \quad (3, -1, 5), \quad (1, -4, 3)$$

在 $\mathbf{R}^3$ 内是线性无关的吗? 对于

$$(5, 4, 3), \quad (3, 3, 2), \quad (8, 1, 3)$$

回答同样的问题.

§23 交错多重线性映射

1. 置换的表示

我们曾经讲过 (§7, 例 4) 对于所有整数 p 用 $\mathfrak{S}_p$ 表示集合 $\{1, 2, \cdots, p\}$ 的置换群. 在这些置换中有所谓对换 (§7, 第 5 小节), 这里对换专指相邻两个整数 i 和 $i+1$

互换, 而其余的整数不变. 在 §7, 第 5 小节曾经看到, 所有的置换 $\sigma \in \mathfrak{S}_p$ 都可以写成对换乘积的形式

$$\sigma = \tau_1 \circ \cdots \circ \tau_r. \tag{1}$$

当然 σ 的分解 (1) 不是唯一的: 举例来说, 如果 τ 是一个对换, 则有 $\tau \circ \tau = e$, 因此

$$\tau = \tau \circ \tau \circ \tau = \tau \circ \tau \circ \tau \circ \tau \circ \tau = \cdots.$$

尽管这样, 在本小节我们将证明当从 σ 的分解 (1) 过渡到另一个时, 整数 r 的奇偶性不改变, 即如果 σ 可以写成偶数 (对应的, 奇数) 个对换的乘积, 那么 σ 的所有对换乘积分解, 都含有偶数 (对应的, 奇数) 个因子.

暂时假定这个结果成立, 那么在分解 (1) 中整数

$$(-1)^r$$

唯一地依赖 σ, 而不依赖 σ 写成对换的乘积的方式, 于是在群 $\mathfrak{S}_p$ 上可以定义一个函数 $\mathfrak{p}$, 其值是 1 和 -1, 使得当 σ 是 r 个对换的乘积时

$$\mathfrak{p}(\sigma) = (-1)^r. \tag{2}$$

显然有

$$\mathfrak{p}(\sigma) = -1, \text{ 如果 } \sigma \text{ 是一个对换}. \tag{3}$$

此外, 如果置换 σ' 和 σ'' 可以分别写成 r 和 s 个对换的乘积, 那么显然 $\sigma' \circ \sigma''$ 可以写成 $r + s$ 个对换的乘积, 故对于任意置换 σ' 和 σ'' 有

$$\mathfrak{p}(\sigma')\mathfrak{p}(\sigma'') = \mathfrak{p}(\sigma' \circ \sigma''). \tag{4}$$

我们注意到 $\{-1, 1\}$ 正是有理整数环 $\mathbf{Z}$ 的可逆元的乘法群 $\mathbf{Z}^*$ (§8, 注 1); 公式 (3) 和 (4) 说明映射

$$\mathfrak{p} : \mathfrak{S}_p \to \mathbf{Z}^* \tag{5}$$

是群的同态, 而且在所有对换上取值为 -1.

反之, 如果构造了一个这样的同态, 那么关系 (1) 给出

$$\mathfrak{p}(\sigma) = \mathfrak{p}(\tau_1) \circ \cdots \circ \mathfrak{p}(\tau_r) = (-1)^r,$$

由于左端仅依赖 σ, 由此推得 r 的奇偶性独立于 σ 的对换乘积的分解 (1).

我们看到, 为了证明在分解 (1) 中的 r 的奇偶性总是同样的 (对给定的置换 σ), 归根结底是要构造一个在每个对换上等于 -1 的一个同态 (5). 我们借助下面的概念实现这一目标.

设 X 是任意一个集合, M 是一个加法群, p 是一个正整数, 并且考虑映射

$$f : \mathrm{X}^p \to \mathrm{M},$$

即 p 个变量 $x_i \in \mathrm{X}$ 在 M 内取值的函数 $f(x_1, \cdots, x_p)$. 称 f 是**反对称的**, 如果对 $i = 1, \cdots, p-1$ 有

$$f(x_1, \cdots, x_{i-1}, x_{i+1}, x_i, x_{i+2}, \cdots, x_p) = -f(x_1, \cdots, x_p), \tag{6}$$

即对于所有对换 τ 有

$$f(x_{\tau(1)}, \cdots, x_{\tau(p)}) = -f(x_1, \cdots, x_p). \tag{7}$$

现在设 f 是一个这样的映射, 我们要指出只要置换 $\sigma \in \mathfrak{S}_p$ 写成 r 个对换的乘积就有

$$f(x_{\sigma(1)}, \cdots, x_{\sigma(p)}) = (-1)^r \cdot f(x_1, \cdots, x_p). \tag{8}$$

情形 $r = 1$ 归结为反对称的定义, 我们要采用关于 r 的归纳推理. 如果 σ 是 r 个对换的乘积, 则有

$$\sigma = \tau \circ \omega,$$

其中 τ 是一个对换, 而 ω 是 $r-1$ 个对换的乘积. 根据归纳假设, 我们对于任意 $y_i \in \mathrm{X}$ 有

$$f(y_{\omega(1)}, \cdots, y_{\omega(p)}) = (-1)^{r-1} \cdot f(y_1, \cdots, y_p). \tag{9}$$

当

$$y_1 = x_{\tau(1)}, \cdots, y_p = x_{\tau(p)}$$

即 $y_i = x_{\tau(i)}$ 时写出这个关系, 并且把 i 换成 $\omega(i)$ 则得

$$y_{\omega(i)} = x_{\tau(\omega(i))} = x_{\sigma(i)}.$$

这时 (9) 写成

$$f(x_{\sigma(1)}, \cdots, x_{\sigma(p)}) = (-1)^{r-1} f(x_{\tau(1)}, \cdots, x_{\tau(p)}).$$

由于 τ 是一个对换, 根据 (7) 这个关系的右端等于关系 (8) 的右端, 因此 (8) 被证明.

(8) 使我们能够证明 r 的奇偶性仅依赖 σ. 由于 (8) 的左端仅依赖 σ, 如果 σ 有 r 个和 s 个对换的两个分解, 则对于所有反对称的 f 有

$$(-1)^r \cdot f(x_1, \cdots, x_p) = (-1)^s \cdot f(x_1, \cdots, x_p).$$

为了得到 $(-1)^r = (-1)^s$, 只需处于这样一种状况, 即可以 "约去" $f(x_1, \cdots, x_p)$, 例如, 如果 f 在 $\mathbf{Z}$ 内取值, 并且存在 $x_1, \cdots, x_p \in \mathrm{X}$ 使得 $f(x_1, \cdots, x_p) \neq 0$. 换句话说,

所有都归结为在适当选择的一个集合 X 上构造一个不恒等于零的在 $\mathbf{Z}$ 内取值的反对称函数.

为此取 $\mathrm{X} = \mathbf{Z}$, 而

$$f(x_1, \cdots, x_p) = \prod_{1 \leqslant i < j \leqslant p} (x_i - x_j). \tag{10}$$

如果 x_i 两两不等, 显然有 $f(x_1, \cdots, x_p) \neq 0$, 剩下要证 f 是反对称的.

假定 x_k 和 x_{k+1} 交换, (10) 右端的因子 $(x_i - x_j)$ 中仅当 $i = k$, $i = k+1$, $j = k$, $j = k+1$ 时被修改, 这些因子对 f 的贡献是

$$\begin{aligned}(x_k - x_{k+1}) \cdot [(x_k - x_{k+2}) \cdots (x_k - x_p)] \cdot [(x_{k+1} - x_{k+2}) \cdots (x_{k+1} - x_p)] \\ \cdot [(x_1 - x_k) \cdots (x_{k-1} - x_k)] \cdot [(x_1 - x_{k+1}) \cdots (x_{k-1} - x_{k+1})];\end{aligned}$$

当 x_k 和 x_{k+1} 交换时, 第一个因子乘以 -1, 而第二个部分乘积和第三个部分乘积互换, 同样第四个部分乘积和第五个部分乘积互换, 所以整个乘积乘以 -1.

于是函数 (10) 确实是反对称的, 比较 (2) 和 (8) 就会发现最终证明了下列结果:

定理 1 对于所有整数 $p \geqslant 1$, 存在唯一的一个同态

$$\mathfrak{p} : \mathfrak{S}_p \to \mathbf{Z}^*$$

使得对于所有对换 σ 有 $\mathfrak{p}(\sigma) = -1$. 如果置换 σ 是 r 个对换的乘积, 则有

$$\mathfrak{p}(\sigma) = (-1)^r.$$

此外, 给定一个集合, 一个加法群 M 和一个反对称映射

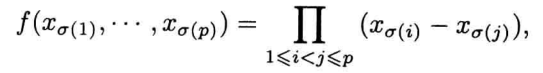
$$f : \mathrm{X}^p \to \mathrm{M},$$

则对于任意 $x_i \in \mathrm{X}$ 和置换 $\sigma \in \mathfrak{S}_p$ 有关系

$$f(x_{\sigma(1)}, \cdots, x_{\sigma(p)}) = \mathfrak{p}(\sigma) \cdot f(x_1, \cdots, x_p).$$

$\mathfrak{p}(\sigma)$ 称为**置换 σ 的符号**, 还经常使用**奇偶性**这个词代替符号这个词, 并且采用记号 ε_σ 代替 $\mathfrak{p}(\sigma)$. 称一个置换是**偶置换**, 如果它的符号是 1; 是**奇置换**, 如果它的符号是 -1. 偶置换组成 $\mathfrak{S}_p$ 的不变子群, 因为它们的集合是同态 $\mathfrak{p}$ 的核 (参见 §7 注 7).

注 1 再看公式 (10), 我们有



$$f(x_{\sigma(1)}, \cdots, x_{\sigma(p)}) = \prod_{1 \leqslant i < j \leqslant p} (x_{\sigma(i)} - x_{\sigma(j)}),$$

这里出现的差 $x_{\sigma(i)} - x_{\sigma(j)}$ 和差 $x_k - x_h (k < h)$ 是相同的, 差别在于这里并不是总有 $\sigma(i) < \sigma(j)$; 我们发现每一个使得

$$i < j, \quad \sigma(i) > \sigma(j)$$

的序偶 (i,j) 在计算 σ 的符号时贡献一个因子 -1. 这些序偶的数目称为 σ 的**逆序数**, 记作 $\mathrm{I}(\sigma)$, 于是有

$$\mathfrak{p}(\sigma)=(-1)^{\mathrm{I}(\sigma)}.$$

例如, 取 $p=6$ 和把 $1,2,3,4,5,6$ 变成

$$2,4,3,6,5,1$$

的置换, 使得 $i<j,\quad \sigma(i)>\sigma(j)$ 的序偶

$$(1,6);(2,3);(2,6);(3,5);(4,5);(4,6);(5,6),$$

故此置换的逆序数是 7, 而符号是 -1.

对于任何 p, 考虑整数 $1,2,\cdots,p$ 的一个**循环置换**, 即一个由

$$\sigma(1)=k,\cdots,\sigma(p-k+1)=p,\sigma(p-k+2)=1,\cdots,\sigma(p)=k-1$$

给定的置换 σ, 其中 k 是 1 和 p 之间的一个整数. 使得 $i<j$, $\sigma(i)>\sigma(j)$ 的序偶是这样的序偶, 对于它们有

$$1\leqslant i\leqslant p-k+1,\quad p-k+2\leqslant j\leqslant p,$$

对于这些序偶, 整数 i 可以取 $p-k+1$ 个不同的值, 而 j 可以取 $k-1$ 个不同的值, 故有

$$\mathrm{I}(\sigma)=(k-1)(p-k+1)=(k-1)(p+1)-k(k-1).$$

由于 $k(k-1)$ 是偶数, 故

$$\mathfrak{p}(\sigma)=(-1)^{(k-1)(p+1)}.$$

例如, 对于 $p=3$ (更一般的, 对于任何奇数), 循环置换是偶置换.

再举一个例子. 设 k,l 是满足 $1\leqslant k<l<p$ 的两个整数, τ 是一个置换, 它使得 k 和 l 交换, l 和 p 之间的其余整数不动, 即

如果 $i\neq k,j$, 则 $\tau(i)=i;\tau(k)=l;\tau(l)=k$, 满足 $i<j$ 和 $\tau(i)>\tau(j)$ 的序偶或者是

$$i=k,\quad k+1\leqslant j\leqslant l,$$

或者是

$$k+1\leqslant i\leqslant l-1, j=l.$$

因此

$$\mathrm{I}(\tau)=(l-k)+(l-1-k)=2(l-k)-1,$$

故得

$$\mathfrak{p}(\tau)=(-1)^{I(\tau)}=-1.$$

于是 τ 是一个奇置换.

2. 多变量函数的反对称化

设 X 是一个集合, M 是一个加法群, $p \geqslant 1$ 是一个整数, 而 f 是从 X^p 到 M 内的一个映射. 称由

$$g(x_1, \cdots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \cdot f(x_{\sigma(1)}, \cdots, x_{\sigma(p)}) \tag{11}$$

定义的从 X^p 到 M 内的映射为 f **的反对称化**, 右端的和是对于所有置换 $\sigma \in \mathfrak{S}_p$ 所取的. 我们就要指出 g 是反对称的, 并且

$$g(x_1, \cdots, x_p) = 0, \quad \text{如果 } x_1, \cdots, x_p \text{ 不是完全不同的}. \tag{12}$$

例 1 如果 $p = 3$, 函数 g 由

$$g(x, y, z) = f(x, y, z) + f(y, z, x) + f(z, x, y) - f(x, z, y) - f(y, x, z) - f(z, y, x)$$

给定, 所宣布的结果差不多是显然的.

为了证实这个结果, 把群 $\mathfrak{S}_p$ 作用在集合 X^p 上, 为此对于 $x = (x_1, \cdots, x_p)$ 令

$$\sigma(x) = (x_{\sigma^{-1}(1)}, \cdots, x_{\sigma^{-1}(p)}).$$

那么公式 (11) 就可以写成

$$g(x) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \cdot f(\sigma^{-1}(x)).$$

设 ω 是任意一个置换, 我们有

$$g(\omega(x)) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \cdot f(\sigma^{-1}(\omega(x))). \tag{13}$$

而对于给定的 ω, 从 $\mathfrak{S}_p$ 到 $\mathfrak{S}_p$ 内的映射 $\sigma \to \omega \circ \sigma$ 是一个双射, 故显然对于定义在群 $\mathfrak{S}_p$ 上的所有函数有

$$\sum_{\sigma \in \mathfrak{S}_p} \varphi(\sigma) = \sum_{\sigma \in \mathfrak{S}_p} \varphi(\omega \circ \sigma);$$

把这个结果运用在

$$\varphi(\sigma) = \mathfrak{p}(\sigma) \cdot f(\sigma^{-1}(\omega(x)))$$

上, 并且观察到

$$\varphi(\omega \circ \sigma) = \mathfrak{p}(\omega \circ \sigma) \cdot f(\sigma^{-1}(\omega^{-1}(\omega(x)))) = \mathfrak{p}(\omega)\mathfrak{p}(\sigma) \cdot f(\sigma^{-1}(x)).$$

我们发现 (13) 还可以写成

$$g(\omega(x)) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\omega)\mathfrak{p}(\sigma) \cdot f(\sigma^{-1}(x)) = \mathfrak{p}(\omega) \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \cdot f(\sigma^{-1}(x));$$

这表明

$$g(\omega(x)) = \mathfrak{p}(\omega) g(x),$$

也就证明了 g 是反对称的.

为了证明 (12), 比如假定对于满足 $i < j$ 的整数 i 和 j 有 $x_i = x_j$; 用 τ 表示如下定义的置换:

$$\tau(k) = k, \quad 如果\ k \neq i, j; \quad \tau(i) = j; \quad \tau(j) = i.$$

显然 $\mathfrak{p}(\tau) = -1$, 对于所考虑的元素 $(x_1, \cdots, x_p) \in \mathrm{X}^p$ 我们有

$$\tau^{-1}(x) = \tau(x) = x. \tag{14}$$

在表达式

$$g(x) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \cdot f(\sigma^{-1}(x))$$

中可以把置换 σ 分成两类: 对于一类有 $\sigma(i) < \sigma(j)$, 而对于另一类有 $\sigma(i) > \sigma(j)$. 用 $\mathfrak{S}'_p$ 和 $\mathfrak{S}''_p$ 表示这样得到的 $\mathfrak{S}_p$ 的两个子集, 它们是不相交的, 并且它们的并集是 $\mathfrak{S}_p$, 我们得到

$$g(x) = \sum_{\sigma \in \mathfrak{S}'_p} \mathfrak{p}(\sigma) \cdot f(\sigma^{-1}(x)) + \sum_{\omega \in \mathfrak{S}''_p} \mathfrak{p}(\omega) \cdot f(\omega^{-1}(x)), \tag{15}$$

而映射 $\sigma \to \tau \circ \sigma$ 是从 $\mathfrak{S}'_p$ 到 $\mathfrak{S}''_p$ 上的双射. 把 (15) 右端的项如此配对; 把第一个和内与 σ 关联的项对应第二个和内与 $\omega = \tau \circ \sigma$ 关联的项, 这两个项的和是

$$\mathfrak{p}(\sigma) \cdot f(\sigma^{-1}(x)) + \mathfrak{p}(\tau \circ \sigma) \cdot f(\sigma^{-1}(\tau^{-1}(x)),$$

由 (14) 得到这个和是

$$\mathfrak{p}(\sigma) \cdot f(\sigma^{-1}(x)) + \mathfrak{p}(\tau)\mathfrak{p}(\sigma) \cdot f(\sigma^{-1}(x)),$$

由于 $\mathfrak{p}(\tau) = -1$, 这个和是零. 如此一来, (15) 中的项两两相消, 这就证明了 (12).

3. 交错多重线性映射

设 X 和 M 是交换环 K 上的模, $p \geqslant 1$ 是一个整数. 称一个 p-重线性映射 $f : \mathrm{X}^p \to \mathrm{M}$ 是**交错的**, 如果只要存在不同的指标 i 和 j, 使得 $x_i = x_j$, 就有 $f(x_1, \cdots, x_p) = 0$. 对于 $p = 1$, 这个概念归结为从 X 到 M 内的线性映射的概念; 对于 $p = 2$ 和 $p = 3$, 就回到前一节的定义.

定理 2 所有交错多重线性映射是反对称的.

为了证明第 1 小节的关系 (6), 我们固定 x_i 和 x_{i+1} 以外的变量的值, 把 f 看作 x_i 和 x_{i+1} 的函数. 由于 f 是多重线性映射, 我们得到 x_i 和 x_{i+1} 的双线性函数, 它还是交错的, 因为当 $x_i = x_{i+1}$ 时 f 变为零. 由此推出 (§22, 第 1 小节, 关系 (2)) 当交换 x_i 和 x_{i+1} 时 $f(x_1, \cdots, x_p)$ 乘以 -1, 这就证明了定理.

根据第 1 小节的结果, 我们看到一个交错 p-重线性映射 $f: \mathrm{X}^p \to \mathrm{M}$ 满足等式

$$f(x_{\sigma(1)}, \cdots, x_{\sigma(p)}) = \mathfrak{p}(\sigma) \cdot f(x_1, \cdots, x_p). \tag{16}$$

对于 $p=2$, 这就是 §22 的关系 (2), 对于 $p=3$, 这就是 §22 的关系 (10).

定理 3 设 f 是从 X^p 到 M 内的一个 p-重线性映射, 那么由

$$g(x_1, \cdots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \cdot f(x_{\sigma(1)}, \cdots, x_{\sigma(p)})$$

给定的 g 是交错 p-重线性映射.

右端的一般项显然是 $x_1, \cdots, x_p$ 的 p-重线性映射, 故 g 也是. 下面证明当向量 $x_1, \cdots, x_p$ 不是两两不同时, $f(x_1, \cdots, x_p) = 0$, 而这是第 2 小节断言 (12) 的一个特殊情形.

例 2 设 $u_1, \cdots, u_p$ 是 X 上的线性型, 那么 p-重线性型

$$\sum \mathfrak{p}(\sigma) u_1(x_{\sigma(1)}) \cdots u_p(x_{\sigma(p)})$$

是交错的: 只需把定理 3 用于在 §21 例 2 定义的型 $u_1 \otimes \cdots \otimes u_p$.

这样得到的交错 p-重线性型称为**线性型** $u_1, \cdots, u_p$ **的外积**, 并且用记号

$$u_1 \wedge \cdots \wedge u_p$$

表示.

定理 4 设 f 是从 X^p 到 M 内的一个交错 p-重线性映射. 只要向量 $a_1, \cdots, a_p$ 是至多 $p-1$ 个向量的线性组合, 就有

$$f(a_1, \cdots, a_p) = 0.$$

事实上, 假定有关系

$$a_i = \sum_{1 \leqslant j \leqslant q} \lambda_{ij} b_j,$$

§21 第 3 小节的公式 (10) 表明

$$f(a_1, \cdots, a_p) = \sum_{j_1, \cdots, j_p} \lambda_{1j_1} \cdots \lambda_{pj_p} f(b_{j_1}, \cdots, b_{j_p}),$$

假定 $q < p$, 那么 1 和 q 之间的 p 个整数 $j_1, \cdots, j_p$ 绝不可能两两互异. 由于 f 是交错的, 对于任意 $j_1, \cdots, j_p$ 必有

$$f(b_{j_1}, \cdots, b_{j_p}) = 0,$$

故得要证明的结果.

推论 1 设 X 和 M 是一个交换域 K 上的线性空间, 而 f 是从 X^p 到 M 内的一个交错 p-重线性映射. 则当 $a_1, \cdots, a_p$ 线性相关时必有

$$f(a_1, \cdots, a_p) = 0.$$

事实上, 如果存在一个非平凡线性关系

$$\lambda_1 a_1 + \cdots + \lambda_p a_p = 0,$$

例如 $\lambda_p \neq 0$, 由于这里 K 是一个域, 由此推出 a_p 是 $a_1, \cdots, a_{p-1}$ 的线性组合. 于是 $a_1, \cdots, a_p$ 是它们当中的 $p-1$ 个的线性组合, 只留下应用定理 4 了.

推论 2 设 X 是具有 r 个向量组成的基的自由的 K-模, 则当 $p \geqslant r+1$ 时从 X^p 到 M 内的一个交错 p-重线性映射是零.

因为对于任意 $x_1, \cdots, x_p$, 可以用 $r < p$ 个向量线性地表示 x_i.

例 3 当 $p > r$ 时在一个 r 维空间上研究 p-重线性映射是无用的, 只需限于研究整数 $p = 1, 2, \cdots, r$ 重的.

如果在特殊情形下 X 是 K $=\mathbf{R}$ 上通常的三维空间, 当 $p \geqslant 4$ 时不存在任何不恒等于零的交错 p-重线性型.

下一小节将使得我们能够把推论 2 精确为这样: 如果 X 具有 r 个向量的基, 则确实存在 X 上不恒等于零的交错的 r-重线性型.

4. 在同构于 K^p 的模上的交错 p-重线性函数

我们要研究从 X^p 到 M 内的交错 p-重线性映射, 这里假定 X 是有限生成的自由模. 在这一小节我们研究 X 同构于 K^p 这一特殊情形, 即 X 具有 p 个向量 $a_1, \cdots, a_p$ 组成的基, 一般情形将是第 7 小节的研究对象.

令

$$x_j = \sum_{i=1}^{p} a_i \xi_{ij},$$

§21 的定理 3 表明(*)

$$f(x_1, \cdots, x_p) = \sum c_{i_1 \cdots i_p} \xi_{i_1 1} \cdots \xi_{i_p p}, \tag{17}$$

(*) 在下面的计算中标量不论写在 M 的元素的左边还是右边都无关紧要, 因为 K 是交换的.

其中

$$c_{i_1\cdots i_p} = f(a_{i_1}, \cdots, a_{i_p}). \tag{18}$$

由于 f 是交错的, 我们有

$$c_{i_1\cdots i_p} = 0, \quad \text{如果 } i_1, \cdots, i_p \text{ 不是两两不同}. \tag{19}$$

于是在 (17) 中可以限于这样的项, p 个整数 $i_1, \cdots, i_p$ 是两两不同的; 但是由于这些整数介于 1 和 p 之间, 它们组成 $1, \cdots, p$ 的一个置换, 即存在唯一的一个置换 $\sigma \in \mathfrak{S}_p$, 使得

$$i_1 = \sigma(1), \cdots, i_p = \sigma(p).$$

这样由于 f 是交错的, 于是就有

$$c_{i_1\cdots i_p} = f(a_{\sigma(1)}, \cdots, a_{\sigma(p)}) = \mathfrak{p}(\sigma) f(a_1, \cdots, a_p).$$

我们发现和式 (17) 可以改写为

$$f(x_1, \cdots, x_p) = f(a_1, \cdots, a_p) \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \xi_{\sigma(1),1} \cdots \xi_{\sigma(p),p}. \tag{20}$$

反之, 从 X^p 到 M 内的所有满足这个关系的映射是交错多重线性的. 为了验证这个事实, 显然只需指出

$$\mathrm{D}(x_1, \cdots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \xi_{\sigma(1),1} \cdots \xi_{\sigma(p),p} \tag{21}$$

是 X^p 上的交错 p-重线性型. 用 $u_1, \cdots, u_p$ 表示模 X 上关于基 $a_1, \cdots, a_p$ 的坐标函数, 则有

$$\xi_{ij} = u_i(x_j),$$

因此有

$$\mathrm{D}(x_1, \cdots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) u_{\sigma(1)}(x_1) \cdots u_{\sigma(p)}(x_p). \tag{22}$$

由于 K 是交换的, 右端的乘积中的 x_i 按照次序 $1, \cdots, p$ 排列, 也可以按照 $\sigma^{-1}(1), \cdots, \sigma^{-1}(p)$ 的次序排列. 我们看到

$$u_{\sigma(1)}(x_1) \cdots u_{\sigma(p)}(x_p) = u_1(x_{\sigma^{-1}(1)}) \cdots u_p(x_{\sigma^{-1}(p)}),$$

故有

$$\mathrm{D}(x_1, \cdots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) u_1(x_{\sigma^{-1}(1)}) \cdots u_p(x_{\sigma^{-1}(p)});$$

但是由于 $\mathfrak{S}_p$ 是一个群, 从 $\mathfrak{S}_p$ 到 $\mathfrak{S}_p$ 内的映射 $\sigma \to \sigma^{-1}$ 是双射. 在这个和中用 σ^{-1} 代替 σ, 只不过相当于更改求和次序, 故有

$$\mathrm{D}(x_1, \cdots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma^{-1}) u_1(x_{\sigma(1)}) \cdots u_p(x_{\sigma(p)});$$

考虑到显然的事实

$$\mathfrak{p}(\sigma^{-1}) = \mathfrak{p}(\sigma)^{-1} = \mathfrak{p}(\sigma),$$

终于得到

$$\mathrm{D}(x_1, \cdots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) u_1(x_{\sigma(1)}) \cdots u_p(x_{\sigma(p)}). \tag{23}$$

这就表明 (例 2) D 正是 X^p 上的线性型 $u_1, \cdots, u_p$ 的外积:

$$\mathrm{D} = u \wedge \cdots \wedge u_p. \tag{24}$$

因此 D 像所断言的那样是 X^p 上的交错 p-重线性型, 并且公式 (20) 刻画了从 X^p 到一个 K-模 M 内的交错 p-重线性映射的特征.

注意有

$$\mathrm{D}(a_1, \cdots, a_p) = 1. \tag{25}$$

事实上, 已经知道

$$u_i(a_j) = \begin{cases} 0, & i \neq j, \\ 1, & i = j, \end{cases}$$

于是如果要根据公式 (22) 计算 $\mathrm{D}(a_1, \cdots, a_p)$, 仅有的可能的非零项是这样的项, 对于它, 有 $\sigma(1) = 1, \cdots, \sigma(p) = p$. 此时有 $u_{\sigma(i)}(a_i) = u_i(a_i) = 1$, 并且显然 $\mathfrak{p}(\sigma) = 1$, 于是得到 (25).

进一步说, 关系 (25) 刻画了 D. 事实上, 根据 (20), 对于 X^p 上的所有交错 p-重线性型我们有

$$f = f(a_1, \cdots, a_p) \cdot \mathrm{D},$$

故关系 $f(a_1, \cdots, a_p) = 1$ 蕴含 $f = \mathrm{D}$.

总之, 我们证明了下列结果:

定理 5 设 X 是交换环 K 上的有限生成的自由模. 设 $(a_1, \cdots, a_p)$ 是 X 的一个基, 则存在唯一的一个 X^p 上的交错 p-重线性型 D, 使得

$$\mathrm{D}(a_1, \cdots, a_p) = 1.$$

对于任意向量

$$x_j = \sum_{i=1}^{p} a_i \xi_{ij} \quad (1 \leqslant j \leqslant p)$$

有

$$\mathrm{D}(x_1, \cdots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \xi_{\sigma(1),1} \cdots \xi_{\sigma(p),p}.$$

最后设 f 是从 X^p 到 K-模 M 内的一个 p-重线性映射, 则对于任意 $x_1, \cdots, x_p \in \mathrm{X}$ 我们有

$$f(x_1, \cdots, x_p) = \mathrm{D}(x_1, \cdots, x_p) f(a_1, \cdots, a_p).$$

这里是这个结果的一个有意思的推论:

推论 设 X 是一个交换环 K 上的有限生成的自由模, 则 X 的所有基有相同数目的元素.

事实上, 设 $(a_1, \cdots, a_p)$ 和 $(b_1, \cdots, b_q)$ 是 X 的两个基, 根据定理 5 存在 X 上的一个交错 q-重线性型 f, 使得

$$f(b_1, \cdots, b_q) = 1,$$

故 f 不恒等于零. 由于 X 具有一个由 p 个向量组成的基, 根据定理 4 的推论应当有 $q \leqslant p$. 但是根据同样的推理还有 $p \leqslant q$, 因此 $p = q$.

5. 向量组、矩阵和自同态的行列式

给定具有基 $(a_1, \cdots, a_p)$ 的一个 K-模 X 和 p 个向量 $x_1, \cdots, x_p$, 称由上一小节关系 (21) 所定义的标量 $\mathrm{D}(x_1, \cdots, x_p)$ 为 $x_1, \cdots, x_p$ **关于基** $(a_1, \cdots, a_p)$ **的行列式**.

另外, 给定元素在 K 内的方阵

$$\mathrm{A} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1p} \\ \vdots & & \vdots \\ \alpha_{p1} & \cdots & \alpha_{pp} \end{pmatrix},$$

称标量

$$\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1p} \\ \vdots & & \vdots \\ \alpha_{p1} & \cdots & \alpha_{pp} \end{vmatrix} = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \alpha_{\sigma(1),1} \cdots \alpha_{\sigma(p),p}$$

为 A **的行列式**, 还记作 $\det(\mathrm{A})$. 显然利用这个定义, 公式 (21) 还可以写作

$$\mathrm{D}(x_1, \cdots, x_p) = \begin{vmatrix} \xi_{11} & \cdots & \xi_{1p} \\ \vdots & & \vdots \\ \xi_{p1} & \cdots & \xi_{pp} \end{vmatrix}.$$

在上一小节看到这个行列式还由公式 (23) 给定, 即它还等于

$$\sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \xi_{1,\sigma(1)} \cdots \xi_{p,\sigma(p)}.$$

这个式子可以由在 (21) 中用 ξ_{ji} 代换 ξ_{ij} 而得到, 故得

定理 6 一个元素在一个交换环内的方阵的行列式等于其转置矩阵的行列式.

这里是另一个重要的结果:

定理 7 设 A 和 B 是元素在一个交换环内的两个 p 阶方阵, 则有

$$\det(\mathrm{AB}) = \det(\mathrm{A})\det(\mathrm{B}).$$

设 X 是一个具有基 $(a_1, \cdots, a_p)$ 的 K-模, 再设 u 和 v 是 X 的同态, 它们关于刚提到的基的矩阵分别是给定的 A 和 B, 那么矩阵 AB 对应于复合映射 $u \circ v$.

设 $\mathrm{D}(x_1, \cdots, x_p)$ 是向量 $x_1, \cdots, x_p$ 关于基 $a_1, \cdots, a_p$ 的行列式. 由

$$\mathrm{D}_u(x_1, \cdots, x_p) = \mathrm{D}(u(x_1), \cdots, u(x_p))$$

定义一个新的从 X^p 到 K 内的映射 D_u. 那么 D_u 仍然是一个交错 p-重线性映射. 首先 D_u 是多重线性的: 事实上举例说, 如果给 $x_2, \cdots, x_p$ 以固定值 $b_2, \cdots, b_p$, 并且令 $c_i = u(b_i)$, 我们得到表达式

$$\mathrm{D}(u(x_1), c_2, \cdots, c_p);$$

作为 x_1 的函数, 它由线性映射 $x_1 \to \mathrm{D}(x_1, c_2, \cdots, c_p)$ 与线性映射 u 复合而得到, 故得到的是 x_1 的线性函数. 如此看来, D_u 是多重线性的, 其次显然它是交错的, 因为关系 $x_i = x_j$ 蕴含 $u(x_i) = u(x_j)$, 因此有 $\mathrm{D}(u(x_1), \cdots, u(x_p)) = 0$.

因为 D_u 是一个交错 p-重线性映射, 定理 5 指出, 对于任意 x_i 有

$$\mathrm{D}_u(x_1, \cdots, x_p) = \mathrm{D}_u(a_1, \cdots, a_p)\mathrm{D}(x_1, \cdots, x_p). \tag{26}$$

如果 $\mathrm{A} = (\alpha_{ij})_{1 \leqslant i,j \leqslant p}$, 则有

$$u(a_j) = \sum_i a_i \alpha_{ij},$$

我们发现

$$\mathrm{D}_u(a_1, \cdots, a_p) = \mathrm{D}(u(a_1), \cdots, (a_p)) = \begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1p} \\ \vdots & & \vdots \\ \alpha_{p1} & \cdots & \alpha_{pp} \end{vmatrix} = \det(\mathrm{A}),$$

于是 (26) 写成

$$\mathrm{D}(u(x_1), \cdots, (x_p)) = \det(\mathrm{A})\mathrm{D}(x_1, \cdots, x_p), \tag{27}$$

同样有

$$\mathrm{D}(v(x_1), \cdots, v(x_p)) = \det(\mathrm{B})\mathrm{D}(x_1, \cdots, x_p),$$

令 $w = u \circ v$, 同样有

$$\mathrm{D}(w(x_1), \cdots, w(x_p)) = \det(\mathrm{AB})\mathrm{D}(x_1, \cdots, x_p). \tag{28}$$

但是

$$\begin{aligned}\mathrm{D}(w(x_1), \cdots, w(x_p)) &= \mathrm{D}(u(v(x_1)), \cdots, u(v(x_p))) \\ &= \det(\mathrm{A})\mathrm{D}(v(x_1), \cdots, v(x_p)) \\ &= \det(\mathrm{A})\det(\mathrm{B})\mathrm{D}(x_1, \cdots, x_p)),\end{aligned}$$

与 (28) 比较即得定理.

公式 (27) 引导出下列概念. 设 u 是 X 的一个同态, 因为交错 p-重线性型 D_u 正比于 D, 存在一个标量, 记作

$$\det(u),$$

使得对于任意向量 $x_i \in \mathrm{X}$ 有

$$\mathrm{D}(u(x_1), \cdots, u(x_p)) = \det(u)\mathrm{D}(x_1, \cdots, x_p).$$

此外根据定理 5, X^p 上的所有交错 p-重线性型都正比于 D, 我们还有

$$f(u(x_1), \cdots, u(x_p)) = \det(u)f(x_1, \cdots, x_p). \tag{29}$$

我们说标量 $\det(u)$ 是**同态 u 的行列式**. 如果 A 是 u 关于 X 的任意一个基的矩阵, 上面的计算 (把 (29) 应用到关于所说的基的行列式) 表明

$$\det(u) = \det(\mathrm{A}).$$

显然, 如果 u 和 v 是 X 上的两个同态, 则有

$$\det(u \circ v) = \det(u)\det(v). \tag{30}$$

由 (29) 显然看出 X 的恒等同态的行列式等于 1.

注 2 前面的内容表明如果 u 是 X 上的一个同态, 那么 u 关于 X 的一个基的行列式不依赖这个基. 还可以通过下面的推理明白这一事实. 设 A 是 u 关于 X 的一个基的矩阵, 那么 u 关于 X 的另一个基的矩阵有形式

$$\mathrm{UAU}^{-1},$$

其中像我们在 §15 所看到的 $\mathrm{U} \in \mathrm{GL}(p, \mathrm{K})$. 于是有

$$\det(\mathrm{UAU}^{-1}) = \det(\mathrm{U})\det(\mathrm{A})\det(\mathrm{U}^{-1}).$$

另一方面有

$$\det(\mathrm{U})\det(\mathrm{U}^{-1}) = \det(\mathrm{I}_p) = 1,$$

故

$$\det(\mathrm{U}^{-1}) = \det(\mathrm{U})^{-1},$$

并且最终得到所断言的

$$\det(\mathrm{UAU}^{-1}) = \det(\mathrm{A}).$$

6. 有限维向量空间基的特征

上一小节的结果蕴含下列定理:

定理 8 设 X 是交换域上的一个向量空间, $a_1, \cdots, a_p$ 是 X 的一个基, 并且

$$x_j = \sum a_i \xi_{ij} \quad (1 \leqslant j \leqslant p)$$

是 X 的 p 个元素, 则下列性质是等价的:

a) 向量 $x_1, \cdots, x_p$ 是线性无关的.

b) 向量 $x_1, \cdots, x_p$ 组成 X 的一个基.

c) 矩阵

$$\begin{pmatrix} \xi_{11} & \cdots & \xi_{1p} \\ \vdots & & \vdots \\ \xi_{p1} & \cdots & \xi_{pp} \end{pmatrix}$$

是可逆的.

d) 我们有

$$\begin{vmatrix} \xi_{11} & \cdots & \xi_{1p} \\ \vdots & & \vdots \\ \xi_{p1} & \cdots & \xi_{pp} \end{vmatrix} \neq 0.$$

a) 和 b) 的等价性来自 §19 定理 10. b) 和 c) 的等价性已经在 §15 定理 1 证明.

用 D 表示关于基 $a_1, \cdots, a_p$ 的行列式, 这是一个交错多重线性型, 条件 d) 改写为 $\mathrm{D}(x_1, \cdots, x_p) \neq 0$. 根据定理 4 的推论 1, 这蕴含 a). 余下要证的是 b) 蕴含 d). 而如果 $x_1, \cdots, x_p$ 组成 X 的一个基, 则存在 (定理 5) X^p 上的一个交错 p-重线性型, 使得

$$f(x_1, \cdots, x_p) \neq 0.$$

由于 X^p 上的所有交错 p-重线性型正比于 D, 更加有

$$\mathrm{D}(x_1, \cdots, x_p) \neq 0,$$

这就是条件 d), 证明完毕.

推论 1 元素在一个交换域内的方阵是可逆的, 必须并且只需它的行列式不是零.

这个结果来自定理 8 的陈述中的 c) 和 d) 的等价性.

推论 2 设 L 是交换域 K 上的一个向量空间, 而 u 是 L 的一个同态, 则以下性质是等价的:

a) u 是双射的.

b) u 是满射的.

c) u 是单射的.

d) 我们有 $\mathrm{Ker}(u)=\{0\}$, 即 $u(x)=0$ 蕴含 $x=0$.

e) u 的行列式不是零.

前四个条件的等价性已经对于不论交换与否的 K 建立 (§19, 定理 13 的推论 1). 此外, u 是双射的, 必须并且只需 (§15, 第 2 小节) 它的关于 L 的一个基的矩阵 A 是可逆的, 根据前一个推论, 即 $\det(\mathrm{A})$ 不是零. 但由于

$$\det(\mathrm{A})=\det(u),$$

故可以看出 a) 和 e) 是等价的.

推论 3 系数在一个交换域 K 内的 n 个未知元 n 个线性齐次方程的方程组

$$\begin{cases}\alpha_{11}\xi_1+\cdots+\alpha_{1n}\xi_n=0,\\ \cdots\cdots\cdots\cdots\\ \alpha_{n1}\xi_1+\cdots+\alpha_{nn}\xi_n=0\end{cases}$$

具有一个非平凡解, 必须并且只需

$$\det((\alpha_{ij}))=0.$$

事实上, 根据 §20 的定理 2 (定理陈述中的 e) 和 f) 之间的等价性), 非平凡解的存在性意味着矩阵 (α_{ij}) 不是可逆的.

推论 4 系数在一个交换域 K 内的 n 个未知元 n 个线性方程的方程组

$$\begin{cases}\alpha_{11}\xi_1+\cdots+\alpha_{1n}\xi_n=\beta_1\\ \cdots\cdots\cdots\cdots\\ \alpha_{n1}\xi_1+\cdots+\alpha_{nn}\xi_n=\beta_n\end{cases}$$

具有唯一解, 必须并且只需

$$\det((\alpha_{ij}))\neq 0.$$

这从 §20 的定理 2 的陈述中的性质 a), d) 和 f) 的等价性得到.

在下一节将会看到行列式理论还提供了 Cramer 方程组的解的明晰公式.



注 3 在下一节将看到推论 1 可以推广到交换环 K, 条件 $\det(\mathrm{A}) \neq 0$ 将换成矩阵 A 是环 K 内的一个可逆元.

注 4 假定 K 是实数域, 而 X 是 p 维实向量空间. 给定 X 的两个基 $(a_1, \cdots, a_p)$ 和 $(b_1, \cdots, b_p)$, 用记号

$$\mathrm{D}(a_1, \cdots, a_p; b_1, \cdots, b_p)$$

表示向量 $b_1, \cdots, b_p$ 关于基 $(a_1, \cdots, a_p)$ 的行列式, 于是它是从基 (a_i) 到基 (b_i) 的过渡矩阵的行列式. 如果有 X 的三个基 $(a_i), (b_i)$ 和 (c_i), 从第一个到第三个的过渡矩阵显然是第一个到第二个的过渡矩阵与第二个到第三个的过渡矩阵的乘积, 于是由定理 7, 我们得到

$$\mathrm{D}(c_1, \cdots, c_p; a_1, \cdots, a_p) = \mathrm{D}(c_1, \cdots, c_p; b_1, \cdots, b_p) \cdot \mathrm{D}(b_1, \cdots, b_p; a_1, \cdots, a_p).$$

由于

$$\mathrm{D}(a_1, \cdots, a_p; a_1, \cdots, a_p) = 1,$$

因此有

$$\mathrm{D}(a_1, \cdots, a_p; b_1, \cdots, b_p) = \mathrm{D}(b_1, \cdots, b_p; a_1, \cdots, a_p)^{-1}.$$

交代了这些, 我们说两个基 $(a_1, \cdots, a_p)$ 和 $(b_1, \cdots, b_p)$ 是**相同方向的**, 如果

$$\mathrm{D}(a_1, \cdots, a_p; b_1, \cdots, b_p) > 0,$$

在相反的情形, 则说它们是**相反方向的**. 我们刚建立的三个关系表明, 两个基有相同方向这个性质在 X 的基的集合中是一个等价关系, 并且基的集合按照这个等价关系刚好分成两类. [为了确信这个事实, 选定一个基 (a_i), 与 (a_i) 有相同方向的基组成第一类; 与 (a_i) 有相反方向的基组成第二类, 因为如果 (b_i) 和 (c_i) 是与 (a_i) 有相反方向的, 则 (b_i) 和 (c_i) 是有相同方向的, 因为两个负数的乘积是正数.]

按照定义, 称这两个等价类中的每一个为 **X 的方向**: X 具有两个可能的方向. 仍然是按照定义, **给 X 指定方向**在于选择了 X 的一个方向, 即基的两个等价类中的一个. 为了给 X 指定方向, 最简单的实施方式是选择 X 的一个基 (a_i), 然后定义我们所选择的方向就是基 (a_i) 所属的那一个类.

给向量空间 X 指定方向之后, 属于 X 上所选择的方向的基称为**顺定向的**或**正定向的**, 其他的基则称为**反定向的**或**负定向的**.

给定了一个任意的向量空间, 没有任何“自然的”或“典范的”或“内在的”手段在 X 里选择一个方向. 即一个“顺定向的基”总假定是任意选择的, 并没有任何绝对的意义. 在物理空间里, 坐标架的“左手”和“右手”法则似乎提供了一个自然的方向的选择, 而“左”和“右”的概念没有任何数学意义.

可以选择典范的方向的仅有的空间是 $\mathbf{R}^n$: 事实上, 宣布和 $\mathbf{R}^n$ 的典范基同一方向的基为顺向的基是自然的. 但是物理的空间仅仅是同构于 $\mathbf{R}^3$, 并不等同于 $\mathbf{R}^3$, 而且不可能不在物理空间上选择基就定义从它到 $\mathbf{R}^n$ 上的同构.

7. 交错多重线性映射: 一般情形

到目前为止, 我们只研究了同构于 K^p 的 K-模上的交错 p-重线性映射. 在一般情形, 我们有类似的更复杂一些的结果:

定理 9 设 X 和 M 是一个交换环 K 上的模, 并且假定 X 是有限生成和自由的, 设 $(a_i)_{1\leqslant i\leqslant n}$ 是 X 的一个基. 从 K^p 到 M 内的 p-重线性映射 f 是交错的, 必须并且只需它关于这个基 (a_i) 的系数

$$c_{i_1\cdots i_p}=f(a_{i_1},\cdots,a_{i_p})$$

满足下列条件:

$$c_{i_1\cdots i_p}=0,\ \text{如果 } i_1,\cdots,i_p \text{ 不是两两不同的;} \tag{31}$$

$$c_{i_{\sigma(1)}\cdots i_{\sigma(p)}}=\mathfrak{p}(\sigma)c_{i_1\cdots i_p}\quad \text{对于所有置换 } \sigma\in\mathfrak{S}_p. \tag{32}$$

如果这些条件满足, 则对于任意向量

$$x_j=\sum_{1\leqslant i\leqslant n}a_i\xi_{ij}\quad (1\leqslant j\leqslant p)$$

有

$$f(x_1,\cdots,x_p)=\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}c_{i_1\cdots i_p}\cdot\begin{vmatrix}\xi_{i_1 1} & \cdots & \xi_{i_1 p}\\ \vdots & & \vdots\\ \xi_{i_p 1} & \cdots & \xi_{i_p p}\end{vmatrix}. \tag{33}$$

令 $a_{i_1}=b_1,\cdots,a_{i_p}=b_p$. 如果指标 $i_1,\cdots,i_p$ 不是两两不同的, 则向量 $b_1,\cdots,b_p$ 中至少有两个相等, 因此, 如果 f 是交错的, 那么有 $f(b_1,\cdots,b_p)=0$, 这就得到 (31). 写出

$$f(b_{\sigma(1)},\cdots,b_{\sigma(p)})=\mathfrak{p}(\sigma)f(b_1,\cdots,b_p),$$

并且注意到

$$b_{\sigma(k)}=a_{i_{\sigma(k)}},$$

则得到关系 (32).

反之, 假定 (31) 和 (32) 满足, 在公式

$$f(x_1,\cdots,x_p)=\sum c_{i_1\cdots i_p}\xi_{i_1 1}\cdots\xi_{i_p p} \tag{34}$$

里, 可以限于对于由两两不同的介于 1 和 n 之间的 p 个整数组成的序列 $i_1,\cdots,i_p$ 求和. 暂时用 S 表示这些序列的集合, 并且设 $\mathrm{S}^+\subset\mathrm{S}$ 是满足

$$i_1<\cdots<i_p \tag{35}$$

的序列 $i_1,\cdots,i_p$ 的集合. 显然所有属于 S 的序列恰可以一种方式从满足 (35) 的一个序列经过适当的置换而得到. 也就是说, 如果令满足 (35) 的每个序列 $(i_1,\cdots,i_p)$ 和每个置换 $\sigma\in\mathfrak{S}_p$ 对应于序列 $(i_{\sigma(1)},\cdots,i_{\sigma(p)})$, 我们就定义了一个从 $\mathrm{S}^+\times\mathfrak{S}_p$ 到 S 上的双射.

于是公式 (34) 可以改写成

$$f(x_1,\cdots,x_p)=\sum_{i_1<\cdots<i_p}\sum_{\sigma\in\mathfrak{S}_p}c_{i_{\sigma(1)}\cdots i_{\sigma(p)}}\xi_{i_{\sigma(1)}1}\cdots\xi_{i_{\sigma(p)}p},$$

考虑到 (32), 我们发现

$$f(x_1,\cdots,x_p)=\sum_{i_1<\cdots<i_p}c_{i_1\cdots i_p}\sum_{\sigma\in\mathfrak{S}_p}\mathfrak{p}(\sigma)\xi_{i_{\sigma(1)}1}\cdots\xi_{i_{\sigma(p)}p}.$$

令 $\alpha_{kh}=\xi_{i_h k}$, 在 $\mathfrak{S}_p$ 上求的部分和改写为

$$\sum\mathfrak{p}(\sigma)\alpha_{\sigma(1),1}\cdots\alpha_{\sigma(p),p},$$

这正是矩阵的行列式

$$\begin{vmatrix}\alpha_{11}&\cdots&\alpha_{1p}\\ \vdots&&\vdots\\ \alpha_{p1}&\cdots&\alpha_{pp}\end{vmatrix}=\begin{vmatrix}\xi_{i_1 1}&\cdots&\xi_{i_1 p}\\ \vdots&&\vdots\\ \xi_{i_p 1}&\cdots&\xi_{i_p p}\end{vmatrix}, \tag{36}$$

这就证明了 (33). 剩下要证明的是 f 是交错的.

用 $u_i(1\leqslant i\leqslant n)$ 表示模 X 关于基 $(a_i)_{1\leqslant i\leqslant n}$ 的坐标函数, 由于按照坐标函数的定义我们有

$$\xi_{ij}=u_i(x_j),$$

故表达式 $\xi_{i_1 1},\cdots,\xi_{i_p p}$ 是 p-重线性型 $u_{i_1}\otimes\cdots\otimes u_{i_p}$ 在向量 $(x_1,\cdots,x_p)$ 上取的值. 因此通过与在第 4 小节 (参见从 (22) 到 (23) 的过渡) 所展开的详细计算类似的计算, 我们看出行列式 (36) 是在第 3 小节例 2 所定义的外积 $u_{i_1}\wedge\cdots\wedge u_{i_p}$ 在向量 $(x_1,\cdots,x_p)$ 取的值. 因此 (33) 改写为

$$f(x_1,\cdots,x_p)=\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}c_{i_1\cdots i_p}u_{i_1\cdots i_p}(x_1,\cdots,x_p), \tag{37}$$

其中

$$u_{i_1\cdots i_p}=u_{i_1}\wedge\cdots\wedge u_{i_p}, \tag{38}$$

由于多重线性型 $u_{i_1}\wedge\cdots\wedge u_{i_p}$ 是交错的 (例 2), 故 f 也是交错的, 这就完成了定理的证明.

当 $\mathrm{M}=\mathrm{K}$ 时, 从前述内容容易推出, 对于 $i_1<\cdots<i_p$, 型 (37) 组成 X^p 上的交错 p-重线性型的模的一个基. 型 (38) 的个数是

$$\binom{n}{p}=\frac{n!}{p!(n-p)!},$$

因为这个二项式系数也是介于 1 和 n 之间的 p 个整数的严格递增序列的个数 (这些序列实际上一一对应到介于 1 和 n 之间的整数集合的 p 个元素的子集).

计算交错 p-重线性型 $u_{i_1\cdots i_p}$ 在基向量上的值是容易的并且是有用的. 结果如下:

$$u_{i_1\cdots i_p}(a_{j_1},\cdots,a_{j_p})=\mathfrak{p}(\sigma),\ \text{如果存在一个置换}\ \sigma\in\mathfrak{S}_p,\ \text{使得}\quad j_1=i_{\sigma(1)},\cdots,j_p=i_{\sigma(p)}; \tag{39}$$

$$u_{i_1\cdots i_p}(a_{j_1},\cdots,a_{j_p})=0,\ \text{在其余的情形}. \tag{40}$$

事实上, (根据 §21 的定理 3) 左端是 $u_{i_1\cdots i_p}(x_1,\cdots,x_p)$ 按向量 $x_1,\cdots,x_p$ 的坐标展开式中 $\xi_{j_1 1}\cdots\xi_{j_p p}$ 的系数, 而此展开式是

$$u_{i_1\cdots i_p}(x_1,\cdots,x_p)=\sum\mathfrak{p}(\sigma)\xi_{i_{\sigma(1)}1}\cdots\xi_{i_{\sigma(p)}p},$$

由此直接得到 (39) 和 (40).

8. 线性无关性的判别法

定理 8 可以推广如下:

定理 10 设 X 是交换域上的一个 n 维向量空间, $(a_1,\cdots,a_n)$ 是 X 的一个基,

$$x_j=\sum_{1\leqslant i\leqslant n}a_i\alpha_{ij}\quad(1\leqslant j\leqslant p)$$

是 X 的元素, 则以下条件是等价的:

a) 向量 $x_1,\cdots,x_p$ 是线性无关的.

b) 存在 X 上的一个交错 p-重线性型 f, 使得

$$f(x_1,\cdots,x_p)\neq 0.$$

c) 可以从矩阵

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1p} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{np} \end{pmatrix}$$

抽取一个其行列式非零的 p 阶方阵.

采用上一小节的记号, 这个矩阵抽取的 p 阶方阵的行列式是标量 $u_{i_1\cdots i_p}(x_1,\cdots,x_p)$. 如果它们当中有一个是非零的, 那么显然条件 b) 将对于适当选取的 $i_1,\cdots,i_p$ 满足. 故 c) 蕴含 b). 此外根据定理 4, b) 蕴含 a).

如果条件 a) 满足, 那么存在 X 的一个基以 $x_1,\cdots,x_p$ 作为开头, 关系 (39) 表明存在一个交错 p-重线性型在 $x_1,\cdots,x_p$ 取非零值. 故 a) 蕴含 b).

剩下要证明的是 b) 蕴含 c). 而公式 (37) 表明如果 $f(x_1,\cdots,x_p)$ 不是零, 至少一个 $u_{i_1\cdots i_p}(x_1,\cdots,x_p)$ 不是零, 这正是性质 c). 定理证明完毕.

注 5 设

$$\mathrm{A} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1p} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{np} \end{pmatrix}$$

是其元素在一个交换域 K 内的矩阵. 根据 §19 的定理 16, A 的秩是使得可以从 A 抽取可逆的即其行列式非零的 r 阶方阵的 r 中的最大者 (定理 8 的推论 1). 上面定理叙述中的条件 c) 表明矩阵 (α_{ij}) 的秩是 p, 那么条件 c) 与条件 a) 的等价性从 §19 的定理 15 得到.

这个结果的重要性在于, 为了计算一个矩阵的秩, 只需检查从它抽出的方阵的行列式, 在实践中这是非常有用的.

此外, 我们发现为了考察 p 个向量 $x_i = \sum \xi_{ij} a_j$ 是线性相关的, 只需写出它的坐标的 $\dbinom{n}{p}$ 个 "代数" 关系, 即

$$\begin{vmatrix} \xi_{i_1 1} & \cdots & \xi_{i_1 p} \\ \vdots & & \vdots \\ \xi_{i_p 1} & \cdots & \xi_{i_p p} \end{vmatrix} = 0, \text{ 任意 } i_1 < \cdots < i_p.$$

例如, 为了表示 $\mathbf{R}^4$ 的三个向量

$$x = (\xi_1, \xi_2, \xi_3, \xi_4),$$
$$y = (\eta_1, \eta_2, \eta_3, \eta_4),$$
$$z = (\zeta_1, \zeta_2, \zeta_3, \zeta_4)$$

是线性相关的, 我们写出

$$\begin{vmatrix} \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \\ \xi_4 & \eta_4 & \zeta_4 \end{vmatrix} = \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_3 & \eta_3 & \zeta_3 \\ \xi_4 & \eta_4 & \zeta_4 \end{vmatrix} = \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_4 & \eta_4 & \zeta_4 \end{vmatrix} = \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix} = 0.$$

9. 线性方程组的相容性条件

行列式理论使得线性方程组的相容性条件 (§19, 定理 5) 置于方便的形式下, 只要基础域是交换的.

设

$$f_i(x) = \alpha_{i1}\xi_1 + \cdots + \alpha_{ip}\xi_p = \beta_i \quad (1 \leqslant i \leqslant n) \tag{41}$$

是系数在 K 内的 p 个未知元 n 个线性方程的方程组. 用 r 表示方程组的秩, 即 K^p 上的线性型 $f_1, \cdots, f_n$ 的族的秩; 或同样, 由 f_j 的系数组成的矩阵 (α_{ij}) 的秩. 于是可以从这个矩阵抽取一个 r 阶可逆的方阵, 即其行列式非零的方阵. 以下假定

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1r} \\ \vdots & & \vdots \\ \alpha_{r1} & \cdots & \alpha_{rr} \end{pmatrix} \neq 0, \tag{42}$$

于是 $f_1, \cdots, f_r$ 是线性无关的, 而 $f_{r+1}, \cdots, f_n$ 是 $f_1, \cdots, f_r$ 的线性组合.

定理 11 假定 (42) 满足, 方程组 (41) 至少具有一个解, 必须并且只需对于所有满足 $r+1 \leqslant j \leqslant n$ 的整数 j 有

$$\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1r} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{r1} & \cdots & \alpha_{rr} & \beta_r \\ \alpha_{j1} & \cdots & \alpha_{jr} & \beta_j \end{vmatrix} = 0.$$

§19 的定理 5 首先指出, 方程组 (41) 具有一个解, 必须并且只需对于所有满足 $r+1 \leqslant j \leqslant n$ 的 j, 方程组

$$\begin{cases} f_1(x) = \beta_1, \\ \cdots\cdots\cdots\cdots \\ f_r(x) = \beta_r, \\ f_j(x) = \beta_j \end{cases}$$

有一个解. 于是可以限于在 $n = r+1$ 的特殊情形下证明定理 11. 以下我们就假定是这种情形.

如果方程组 (41) 有解, 那么我们通过解前 r 个方程构成的方程组得到这些解, 而由于假设 (42) 表明线性方程组

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{1r}\xi_r = \beta_1, \\ \cdots\cdots\cdots\cdots \\ \alpha_{r1}\xi_1 + \cdots + \alpha_{rr}\xi_r = \beta_r \end{cases}$$

是 Cramer 方程组 (§20, 定理 2 或本节定理 8 的推论 4), 我们发现可以把任意值赋予出现在 (41) 中的未知元 $\xi_{r+1}, \cdots, \xi_n$ (§20, 第 5 小节), 特别的, 如果方程组 (41) 具有一个解, 则它具有一个这样的解:

$$\xi_{r+1} = \cdots = \xi_p = 0,$$

而逆命题显然是平凡的. 在我们所感兴趣的 $n = r + 1$ 这种情形, 所有事情归结为说明秩为 r 的、r 个未知元 $r + 1$ 个方程的方程组

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{r1}\xi_r = \beta_1, \\ \cdots\cdots\cdots\cdots \\ \alpha_{1,r+1}\xi_1 + \cdots + \alpha_{r,r+1}\xi_r = \beta_{r+1} \end{cases} \tag{43}$$

至少有一个解, 并且指出它有解必须并且只需

$$\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1r} & \beta_1, \\ \vdots & & \vdots & \vdots \\ \alpha_{r1} & \cdots & \alpha_{rr} & \beta_r \\ \alpha_{r+1,1} & \cdots & \alpha_{r+1,r} & \beta_{r+1} \end{vmatrix} = 0. \tag{44}$$

而条件 (44) 还表示方程组

$$\begin{cases} \alpha_{11}\eta_1 + \cdots + \alpha_{1r}\eta_r + \beta_1\eta_{r+1} = 0, \\ \cdots\cdots\cdots\cdots \\ \alpha_{r+1,1}\eta_1 + \cdots + \alpha_{r+1,r}\eta_r + \beta_{r+1}\eta_{r+1} = 0 \end{cases} \tag{45}$$

有一个非平凡解 (定理 8 的推论 3). 于是我们只需证明在假设 (42) 之下, (43) 一个*解的存在性等价于 (45) 非平凡解的存在性*.

首先, 显然第一个性质蕴含第二个性质, 因为如果 $(\xi_1, \cdots, \xi_r)$ 是 (43) 的一个解, 那么 $(\xi_1, \cdots, \xi_r, -1)$ 是 (45) 的非平凡解.

再者, 反之考虑 (45) 的一个非平凡解 $(\eta_1, \cdots, \eta_{r+1})$, 那么有

$$\eta_{r+1} \neq 0, \tag{46}$$

因为否则 $\eta_{r+1}=0$, $(\eta_1,\cdots,\eta_r)$ 将是与 (43) 相伴的齐次方程组 (43) 的一个非平凡解, 更是

$$\begin{cases}\alpha_{11}\eta_1+\cdots+\alpha_{1r}\eta_r=0,\\ \qquad\cdots\cdots\cdots\cdots\\ \alpha_{r1}\eta_1+\cdots+\alpha_{rr}\eta_r=0\end{cases}$$

的非平凡解, 而这与假设 (42) 和定理 8 的推论 3 相矛盾. 于是必有 $\eta_{r+1}\neq 0$, 这就允许用 η_{r+1} 除方程 (45), 并且发现标量

$$\xi_i=-\eta_i/\eta_{r+1}\quad(1\leqslant i\leqslant r)$$

满足 (43), 这就完成了证明.

例 4 取 K =**R**, 考虑三个未知元的线性方程组

$$\begin{cases}x+2y+3z=a,\\ 4x+5y+6z=b,\\ 7x+8y+9z=c,\end{cases}$$

它的行列式

$$\begin{vmatrix}1&2&3\\4&5&6\\7&8&9\end{vmatrix}=1\cdot\begin{vmatrix}5&6\\8&9\end{vmatrix}-2\cdot\begin{vmatrix}4&6\\7&9\end{vmatrix}+3\cdot\begin{vmatrix}4&5\\7&8\end{vmatrix}$$

直接看出是零, 故所考虑的方程组不是一个 Cramer 方程组. 由于

$$\begin{vmatrix}1&2\\4&5\end{vmatrix}=-3$$

不是零, 方程组的秩是 2; 相容性条件仅有一个, 即

$$\begin{vmatrix}1&2&a\\4&5&b\\7&8&c\end{vmatrix}=0,$$

容易看出这个条件可以写成

$$3(2b-a-c)=0,$$

或写成(*)

$$2b=a+c.$$

(*) 对于那些熟悉域的特征 (§30, 第 6 小节) 的读者来说, 应当对 K 是任意的交换域时所考虑的方程组进行更完全的研究; 举例说, 显然文中的推理当域的特征是 3 时就失效了.

由于显然有

$$(x+2y+3z)+(7x+8y+9z)=2(4+5y+6z),$$

所发现的条件的必要性可以事先预见得到.



注 6 读者比较定理 11 和 §19 的习题 23 会有益处.

§23 习题

1. 设 K 是一个交换域. 证明满足条件

$$\det(\mathrm{U})=1$$

的矩阵 $\mathrm{U}\in \mathrm{M}_n(\mathrm{K})$ 组成 $\mathrm{GL}(n,\mathrm{K})$ 的一个子群, 通常将这个子群记作 $\mathrm{SL}(n,\mathrm{K})$, 并且将它称为环 K 上的 n 个变量的**特殊线性群**.

2. 元素在交换域内的幂零矩阵的行列式是零.

3. 设 U 是元素为整数的方阵, 其行列式非零. 证明只有能整除 U^{-1} 的 (有理数) 元素的分母的素数是 U 的行列式的素数因子.

4. 设 U 是元素在一个交换域内的正交方阵 (即 ${}^t\mathrm{U}\cdot\mathrm{U}=1$). 证明 $\det(\mathrm{U})=1$ 或 -1.

5. 求置换

$$\begin{pmatrix}1 & 2 & 3 & \cdots & n & n+1 & n+2 & n+3 & \cdots & 2n\\ 1 & 3 & 5 & \cdots & 2n-1 & 2 & 4 & 6 & \cdots & 2n\end{pmatrix}$$

的逆序数.

6. 整数 $1,2,\cdots,n$ 的所有置换中哪个有最大的逆序数?

7. 证明对于每一个满足条件 $0\leqslant k\leqslant \binom{n}{2}$ 的整数 k, 存在整数 $1,2,\cdots,n$ 的一个逆序数为 k 的置换.

8. 考虑六阶行列式, 其元素用 $a_{ij}(1\leqslant i,j\leqslant 6)$ 表示. 在这个行列式展开式中乘积

$$a_{61}a_{23}a_{45}a_{36}a_{12}a_{54}$$

应该冠以什么符号?

9. 在 $\{1,2,\cdots,n\}$ 的置换群 $\mathfrak{S}_n$ 中, 用 $\mathfrak{A}_n$ 表示偶置换的集合.

a) 证明 $\mathfrak{A}_n$ 是 $\mathfrak{S}_n$ 的不变子群 (称为 n 个对象的**交错群**), 并且商群 $\mathfrak{S}_n/\mathfrak{A}_n$ 同构于 $\mathbf{Z}/2\mathbf{Z}$.

b) 对于 $3\leqslant i\leqslant n$ (假定 $n\geqslant 3$), 用 s_i 表示置换

$$\begin{pmatrix}1 & 2 & 3 & \cdots & i-1 & i & i+1 & \cdots & n\\ i & 1 & 3 & \cdots & i-1 & 2 & i+1 & \cdots & n\end{pmatrix},$$

证明 s_i 生成 $\mathfrak{A}_n$.

c) 证明对于 $n\geqslant 5$, $\mathfrak{A}_n$ 的仅有的不变子群是 $\mathfrak{A}_n$ 自己和缩减为恒等置换的子群 (这表示对于 $n\geqslant 5$, $\mathfrak{A}_n$ 是**单群**). 对于 $n=2,3,4$, $\mathfrak{A}_n$ 的不变子群是什么?

d) 证明对于 $n\neq 4$, $\mathfrak{S}_n$ 仅有的非平凡不变子群是 $\mathfrak{A}_n$.

¶10. 设 $\mathrm{A} = (a_{ij})$ 是元素在一个交换域 K 内的 n 阶可逆方阵. 求元素在 K 中的 n 阶方阵 X 和 Y, 使得它们满足关系

$$\mathrm{A} = \mathrm{XY},$$

并且有形式

$$\mathrm{X} = \begin{pmatrix} * & 0 & 0 & \cdots & 0 \\ * & * & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ * & * & * & \cdots & * \end{pmatrix}, \quad \mathrm{Y} = \begin{pmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & * \end{pmatrix}$$

(其中的符号 $*$ 表示 K 的任意元素). 证明 X 和 Y 存在, 必须并且只需

$$\begin{vmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pp} \end{vmatrix} \neq 0 \quad 对于\ 1 \leqslant p \leqslant n-1.$$

在这种情形, 可以要求 X (或 Y) 的对角线元素都是 1, 而且这个条件完全确定了 X 和 Y.

11. 设 K 是一个交换环. 称从 K 到 K 内的所有映射 D 为 K 的**导子**, 如果 D 满足条件: 对于任意 $x, y \in \mathrm{K}$ 有

$$\mathrm{D}(x+y) = \mathrm{D}(x) + \mathrm{D}(y), \quad \mathrm{D}(xy) = \mathrm{D}(x) \cdot y + x \cdot \mathrm{D}(y).$$

设 D 为 K 的一个导子, 而 $\mathrm{A} = (a_{ij})_{1 \leqslant i,j \leqslant n}$ 是元素在 K 内的 n 阶方阵. 对于每个满足 $1 \leqslant i \leqslant n$ 的 i, 用 A_i 表示由 A 把 D 应用到 A 的第 i 列的每一个元素得到的矩阵. 证明

$$\mathrm{D}(\det(\mathrm{A})) = \det(\mathrm{A}_1) + \cdots + \det(\mathrm{A}_n)$$

(行列式求导法则).

¶¶12. 设 M 是交换环 K 上的一个模, f 是 M 上的一个交错 p-重线性型, 而 g 是 M 上的一个交错 q-重线性型. 定义一个映射

$$h : \mathrm{M}^{p+q} \to \mathrm{K} \quad (\mathrm{K}\ 是基础环)$$

如下:

$$h(x_1, \cdots, x_{p+q}) = \sum_{\substack{s \in \mathfrak{S}_{p+q} \\ s(1) < \cdots < s(p) \\ s(p+1) < \cdots < s(p+q)}} \mathfrak{p}(s) \cdot f(x_{s(1)}, \cdots, x_{s(p)}) \cdot g(x_{s(p+1)}, \cdots, x_{s(p+q)}), \qquad (*)$$

其中求和是取自整数 $1, \cdots, p+q$ 的所有这样的置换 s: 它们保持这些整数中的前 p 个的次序和后 q 个的次序.

a) 证明 h 是 M 上的交错 $p+q$-重线性型, 称为型 f 和 g 的**外积**, 并且用记号

$$h = f \wedge g$$

表示.

b) 证明

$$g \wedge f = (-1)^{pq} f \wedge g.$$

c) 证明如果 f, g, h 是 M 上的三个交错多重线性型, 则有

$$f \wedge (g \wedge h) = (f \wedge g) \wedge h$$

(外积的 "结合性").

d) 设 $u_1, \cdots, u_p, v_1, \cdots, v_q$ 是 M 上的线性型. 在公式 (*) 中取

$$f = u_1 \wedge \cdots \wedge u_p, \quad g = v_1 \wedge \cdots \wedge v_q$$

(参见 §23, 第 3 小节, 例 2). 证明

$$h = u_1 \wedge \cdots \wedge u_p \wedge v_1 \wedge \cdots \wedge v_q.$$

13. (行列式乘法定理的推广) 设

$$\mathrm{A} = (a_{ij})_{1 \leqslant i \leqslant p, 1 \leqslant j \leqslant q}$$

是元素在 K 内的一个矩阵, 取一个整数 r, 满足条件 $1 \leqslant r \leqslant p$ 和 $1 \leqslant r \leqslant q$. 给定集合 $\{1, \cdots, p\}$ 的 r 个元素的一个子集 I 和集合 $\{1, \cdots, q\}$ 的 r 个元素的一个子集 J, 用 a_{IJ} 表示 r 阶方阵, 它由 $i \in \mathrm{I}$ 且 $j \in \mathrm{J}$ 的 a_{ij} 组成; 最后用

$$\Lambda^r(\mathrm{A})$$

表示矩阵

$$(\det(a_{\mathrm{IJ}}))_{\mathrm{I} \subset \{1, \cdots, p\}, \mathrm{J} \subset \{1, \cdots, q\}, \mathrm{Card(I)} = \mathrm{Card(J)} = r}.$$

可以对于 $\{1, \cdots, p\}$ 的 r 个元素的子集, 举例说, 按照字典排序法排序, 即约定

$$\{i_1, \cdots, i_r\}, \quad \text{其中 } i_1 < \cdots < i_r,$$

先于

$$\{j_1, \cdots, j_r\}, \quad \text{其中 } j_1 < \cdots < j_r,$$

如果存在一个整数 $h (1 \leqslant h \leqslant r)$ 使得

$$i_1 = j_1, \cdots, i_{h-1} = j_{h-1}, i_h < j_h,$$

(参见在字典里的词的编号方法) 标量 $\det(a_{\mathrm{IJ}})$ 称为矩阵 A 的**子式**.

证明如果 A 和 B 是元素在 K 内的矩阵, 并且乘积 AB 有意义, 则有

$$\Lambda^r(\mathrm{AB}) = \Lambda^r(\mathrm{A}) \Lambda^r(\mathrm{B}).$$

¶14. 设 A 和 B 是 p 行 q 列的元素在交换环 K 内的两个矩阵. 称 A 和 B 是**等价的** (在基础环 K 上), 如果存在

$$\mathrm{U} \in \mathrm{GL}(p, \mathrm{K}), \quad \mathrm{V} \in \mathrm{GL}(q, \mathrm{K}),$$

使得 $\mathrm{B} = \mathrm{UAV}$.

如果是这样, 证明对于所有 $r \leqslant \min(p,q)$, 由 A 的 r 阶子式生成的 K 的理想等于由 B 的 r 阶子式生成的 K 的理想. [注意: 如果 K 是主理想整环, 则其逆亦真; 参见 §31, 习题 11, e).]

¶15. 设 K 是交换环, 而 $A \in M_p(K), B \in M_q(K)$. 考虑 [§21, 习题 4, f)] 矩阵 $A \otimes B \in M_{pq}(K)$. 证明

$$\det(A \otimes B) = \det(A)^q \det(B)^p.$$

¶¶16. 设 A 是元素在交换环 K 内的 n 阶方阵. 对于 $1 \leqslant r \leqslant n$, 考虑习题 13 的矩阵 $\Lambda^r(A)$. 证明

$$\det(\Lambda^r(A)) = \det(A)^{\binom{n-1}{r-1}}.$$

¶¶17. 设 M 是交换环上的一个有限生成的自由模, 而 u 是 M 的一个自同构. 通过 u 的行列式计算 $T_q^p(M)$ 的同构 $T_q^p(u)$ (§21, 习题 1) 的行列式.

¶18. 设 A 是元素在一个交换环 K 内的奇数阶方阵. 假定 A 是反对称的, 即 ${}^tA = -A$. 证明 $\det(A) = 0$ (利用 §22 的习题 17).

§24 行列式

1. 行列式的基本性质

给定一个元素在一个交换环 K 内的方阵

$$X = \begin{pmatrix} \xi_{11} & \cdots & \xi_{1n} \\ \vdots & & \vdots \\ \xi_{n1} & \cdots & \xi_{nn} \end{pmatrix},$$

我们在 §23 第 5 小节定义它的行列式是标量

$$\det(X) = \sum_{\sigma \in \mathfrak{S}_n} \mathfrak{p}(\sigma) \xi_{\sigma(1),1} \cdots \xi_{\sigma(n),n}.$$

如果在 K^n 里引进同态 u, 它 (关于 K^n 的典范基) 的矩阵是 X, 而向量

$$x_j = u(e_j) = (\xi_{1j}, \cdots, \xi_{nj})$$

用 X 的列向量表示, 则有

$$\det(X) = \det(u) = D(x_1, \cdots, x_n),$$

这里 $D(x_1, \cdots, x_n)$ 表示向量 x_i 关于 K^n 的典范基向量 $e_1, \cdots, e_n$ 的行列式.

由此显然得到 X 的行列式是 X 的列向量的交错多重线性型. 还得到在实际中重要的计算法则:

a) 两列相等的行列式等于零.

因为交错多重线性型当它含有两个相等的变量时为零.

b) 如果行列式的列经受一个置换 σ, 则行列式的值乘以 σ 的符号, 其特殊情形是, 当交换两列时, 行列式的值乘以 -1.

这个性质来源于对于所有交错多重线性函数都有效的等式

$$f(x_{\sigma(1)}, \cdots, x_{\sigma(n)}) = \mathfrak{p}(\sigma) f(x_1, \cdots, x_n).$$

例如:

$$\begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix} = -\begin{vmatrix} b & a & c \\ b' & a' & c' \\ b'' & a'' & c'' \end{vmatrix} = \begin{vmatrix} c & a & b \\ c' & a' & b' \\ c'' & a'' & b'' \end{vmatrix}.$$

c) 如果在一个行列式里, 一个给定的列的每个元素都乘以同一个标量 λ, 则行列式乘以 λ.

d) 假定对于一个给定的整数 j, 矩阵 X 的第 j 列的形式是

$$\xi_{ij} = \xi'_{ij} + \xi''_{ij} \quad (1 \leqslant i \leqslant n).$$

设 X' (对应的, X'') 是通过对于 $1 \leqslant i \leqslant n$ 用 ξ'_{ij} (对应的, ξ''_{ij}) 代替 ξ_{ij} 从 X 得到的矩阵, 则

$$\det(\mathrm{X}) = \det(\mathrm{X}') + \det(\mathrm{X}'').$$

性质 c) 和 d) 分别对应于 §21 的关系 (3) 和 (4).

例如我们有

$$\begin{vmatrix} a & u+2v & c \\ a' & u'+2v' & c' \\ a'' & u''+2v'' & c'' \end{vmatrix} = \begin{vmatrix} a & u & c \\ a' & u' & c' \\ a'' & u'' & c'' \end{vmatrix} + 2 \cdot \begin{vmatrix} a & v & c \\ a' & v' & c' \\ a'' & v'' & c'' \end{vmatrix}.$$

结合 a), c) 和 d) 得到

e) 行列式的一个列加上其他列的一个线性组合, 其值不变.

例如考虑行列式

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix};$$

先从第三列减去第二列, 再从第二列减去第一列, 得到行列式

$$\begin{vmatrix} 1 & 1 & 1 \\ 4 & 1 & 1 \\ 7 & 1 & 1 \end{vmatrix},$$

由于它的两列相等, 故等于零.

另外, §23 定理 6 的关系

$$\det({}^t\mathrm{X}) = \det(\mathrm{X})$$

表明

f) 当交换行列式的行和列时, 其值不变.

由此得到一个行列式是既是它的列的也是它的行的交错多重线性函数. 因此有

g) 法则 a), b), c), d), e) 当把 "列" 换成 "行" 时保持有效.

2. 行列式按一行或一列的展开

设 X 是一个 K-模, 它有一个由 n 个元素组成的基 $(a_1, \cdots, a_n)$, 考虑 X 上的交错 $(n-1)$-重线性型. §23 的定理 9 表明如果

$$x_j = \sum a_i \xi_{ij},$$

则有

$$f(x_1, \cdots, x_{n-1}) = \sum_{1 \leqslant i_1 < \cdots < i_{n-1} \leqslant n} \gamma_{i_1 \cdots i_{n-1}} \cdot \begin{vmatrix} \xi_{i_1,1} & \cdots & \xi_{i_1,n-1} \\ \vdots & & \vdots \\ \xi_{i_{n-1},1} & \cdots & \xi_{i_{n-1},n-1} \end{vmatrix},$$

其中

$$\gamma_{i_1 \cdots i_{n-1}} = f(a_{i_1}, \cdots, a_{i_{n-1}}).$$

而 $n-1$ 个介于 1 和 n 之间的整数 $i_1, \cdots, i_{n-1}$ 是两两不同的, 出现在前面和式中的 $i_1 < \cdots < i_{n-1}$ 共有 n 个, 它们是

$$(2, \cdots, n); (1, 3, \cdots, n); \cdots; (1, \cdots, n-1),$$

即这是序列

$$(1, \cdots, i-1, i+1, \cdots, n), \text{ 其中 } 1 \leqslant i \leqslant n.$$

于是前面的公式还可以写成

$$f(x_1, \cdots, x_{n-1}) = \sum_{1 \leqslant i \leqslant n} \gamma_i \mathrm{D}_i(x_1, \cdots, x_{n-1}), \tag{1}$$

其中的记号是

$$\mathrm{D}_i(x_1, \cdots, x_{n-1}) = \begin{vmatrix} \xi_{11} & \cdots & \xi_{1,n-1} \\ \vdots & & \vdots \\ \xi_{i-1,1} & \cdots & \xi_{i-1,n-1} \\ \xi_{i+1,1} & \cdots & \xi_{i+1,n-1} \\ \vdots & & \vdots \\ \xi_{n1} & \cdots & \xi_{n,n-1} \end{vmatrix}, \tag{2}$$

和

$$\gamma_i = f(a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_n). \tag{3}$$

我们注意 $\mathrm{D}_i(x_1, \cdots, x_{n-1})$ 是从由向量 x_i 的分量所组成的矩阵

$$\begin{pmatrix} \xi_{11} & \cdots & \xi_{1,n-1} \\ \vdots & & \vdots \\ \xi_{n1} & \cdots & \xi_{n,n-1} \end{pmatrix} \tag{4}$$

去掉第 i 行所得到的矩阵的行列式.

这个事实建立之后, 取 X 为模 K^n, K^n 的典范基作为 X 的基, 并且由表达式

$$f(x_1, \cdots, x_{n-1}) = \mathrm{D}(x_1, \cdots, x_{j-1}, u, x_j, \cdots, x_{n-1}) \tag{5}$$

定义 f, 其中 u 是 K^n 的一个固定元素, 而 D 是关于典范基的行列式. 显然, 由于 D 是交错 n-重线性的, f 必然是 K^n 上的一个交错 $(n-1)$-重线性型. 令

$$u = \sum_{1 \leqslant i \leqslant n} \alpha_i e_i,$$

根据第 1 小节有

$$f(x_1, \cdots, x_{n-1}) = \begin{vmatrix} \xi_{11} & \cdots & \xi_{1,j-1} & \alpha_1 & \xi_{1j} & \cdots & \xi_{1,n-1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \xi_{n1} & \cdots & \xi_{n,j-1} & \alpha_n & \xi_{nj} & \cdots & \xi_{n,n-1} \end{vmatrix}. \tag{6}$$

另外表达式 (2) 是向量 $x_1, \cdots, x_{n-1}$ 的指标不等于 i 的坐标组成的 $n-1$ 阶行列式. 还需要计算

$$\begin{aligned} \gamma_j &= f(e_1, \cdots, e_{i-1}, e_{i+1}, \cdots, e_n) \\ &= \begin{cases} \mathrm{D}(e_1, \cdots, e_{i-1}, e_{i+1}, \cdots, e_{j-1}, u, e_j, \cdots, e_n), & \text{如果 } i \leqslant j, \\ \mathrm{D}(e_1, \cdots, e_{j-1}, u, e_j, \cdots, e_{i-1}, e_{i+1}, \cdots, e_n), & \text{如果 } i > j; \end{cases} \end{aligned}$$

由于向量 u 是向量 $\alpha_i e_i$ 与出现在要计算的行列式中的向量

$$e_1, \cdots, e_{i-1}, e_{i+1}, \cdots, e_n$$

的线性组合之和, 故可以用 $\alpha_i e_i$ 代替 u, 因此得到

$$\gamma_i = \begin{cases} \alpha_i \cdot \mathrm{D}(e_1, \cdots, e_{i-1}, e_{i+1}, \cdots, e_{j-1}, e_i, e_j, \cdots, e_n), & \text{如果 } i \leqslant j, \\ \alpha_j \cdot \mathrm{D}(e_1, \cdots, e_{j-1}, e_i, e_j, \cdots, e_{i-1}, e_{i+1}, \cdots, e_n), & \text{如果 } i > j. \end{cases}$$

由此立即得到

$$\gamma_i = (-1)^{i+j} \alpha_i \cdot \mathrm{D}(e_1, \cdots, e_n) = (-1)^{i+j} \alpha_i.$$

把这些结果代入关系 (1) 即得等式

$$f(x_1,\cdots,x_{n-1})=\sum_{j=1}^{n}(-1)^{i+j}\alpha_i\cdot\begin{vmatrix}\xi_{11} & \cdots & \xi_{1,n-1}\\ \vdots & & \vdots\\ \xi_{i-1,1} & \cdots & \xi_{i-1,n-1}\\ \xi_{i+1,1} & \cdots & \xi_{i+1,n-1}\\ \vdots & & \vdots\\ \xi_{n,1} & \cdots & \xi_{n,n-1}\end{vmatrix}.$$

考虑到 (6), 并且采用更对称的符号, 我们得到下列结果:

定理 1 *设 $X=(\xi_{ij})_{1\leqslant i,j\leqslant n}$ 是一个元素在交换环 K 内的 n 阶方阵. 用 X_{ij} 表示在 X 内去掉第 i 行和第 j 列的矩阵. 则对于所有满足 $1\leqslant j\leqslant n$ 的 j 都有*

$$\det(X)=\sum_{1\leqslant i\leqslant n}(-1)^{i+j}\xi_{ij}\cdot\det(X_{ij}). \tag{7}$$

由于标量 $\det(X_{ij})$ 与 X 的第 j 列无关, 公式 (7) 使得 X 的行列式是出现在 X 的第 j 列的元素的线性函数的事实 (由多重线性型的定义这是显然的) 明晰地表示出来. 基于这个理由, 我们说 (7) 是 det(X) **按照 X 的第 j 列的展开式**.

由于 $\det({}^tX)=\det(X)$, 必定还有公式

$$\det(X)=\sum_{1\leqslant j\leqslant n}(-1)^{i+j}\xi_{ij}\cdot\det(X_{ij}), \tag{8}$$

称为 det(X) **按照 X 的第 i 行的展开式.**

例 1 取 X 是一个**三角矩阵**, 即形式为

$$X=\begin{pmatrix}\alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n}\\ 0 & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n}\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & \alpha_{nn}\end{pmatrix}$$

的矩阵. 按照它的第一列展开得到

$$\det(X)=\alpha_{11}\cdot\begin{vmatrix}\alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n}\\ 0 & \alpha_{33} & \cdots & \alpha_{3n}\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & \alpha_{nn}\end{vmatrix}.$$

由此利用关于 n 的数学归纳法得到

$$\det(X)=\alpha_{11}\alpha_{22}\cdots\alpha_{nn},$$

这是 X 的对角元素的乘积.

这个公式是下列公式的特殊情形. 设 $n_1, \cdots, n_p$ 是正整数, 并考虑形式为

$$
\mathrm{X} = \begin{pmatrix} \mathrm{A}_{11} & \mathrm{A}_{12} & \cdots & \mathrm{A}_{1p} \\ 0 & \mathrm{A}_{22} & \cdots & \mathrm{A}_{2p} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \mathrm{A}_{pp} \end{pmatrix}
$$

的方阵, 其中 A_{ij} 为 n_i 行 n_j 列的矩阵. 则有

$$
\det(\mathrm{X}) = \det(\mathrm{A}_{11}) \det(\mathrm{A}_{22}) \cdots \det(\mathrm{A}_{pp}).
$$

为了证明这个结果, 只需对于 p 进行归纳推理, 先证明它当 $p = 2$ 时成立, 即指出如果 A 是 p 阶方阵, D 是 q 阶方阵, B 是 p 行 q 列矩阵, 则有

$$
\det \begin{pmatrix} \mathrm{A} & \mathrm{B} \\ 0 & \mathrm{D} \end{pmatrix} = \det(\mathrm{A}) \cdot \det(\mathrm{D}).
$$

为此, 让我们考虑 K^{p+q}, 并且设 u 是其关于典范基 $(e_1, \cdots, e_{p+q})$ 的矩阵为 X 的 K^{p+q} 的自同态, 设 E 是由 $(e_1, \cdots, e_p)$ 生成的子空间, 而 F 是由 $(e_{p+1}, \cdots, e_{p+q})$ 生成的子空间, 这样就有

$$
\mathrm{K}^{p+q} = \mathrm{E} \oplus \mathrm{F};
$$

设 v 是关于基 $(e_1, \cdots, e_p)$ 以 A 为其矩阵的 E 的同态, 而 w 是关于基 $(e_{p+1}, \cdots, e_{p+q})$ 以 D 为其矩阵的 F 的同态, 而 $\mathrm{D}(x_1, \cdots, x_{p+q})$ 是变向量 $x_1, \cdots, x_{p+q}$ 关于基 $(e_1, \cdots, e_{p+q})$ 的行列式, 则有

$$
\det(\mathrm{X}) = \mathrm{D}(u(e_1), \cdots, u(e_{p+q})).
$$

显然有

$$
u(e_1) = v(e_1), \cdots, u(e_p) = v(e_p),
$$
$$
\det(\mathrm{X}) = \mathrm{D}(v(e_1), \cdots, v(e_p), b_{p+1}, \cdots, b_{p+q}),
$$

其中我们临时记

$$
b_{p+j} = u(e_{p+j}) \quad (1 \leqslant j \leqslant q).
$$

如果 $b_{p+1}, \cdots, b_{p+q}$ 是给定的, 那么显然表达式

$$
\mathrm{D}(x_1, \cdots, x_p, b_{p+1}, \cdots, b_{p+q})
$$

当 x_i 在 E 内变化时是 E 上的交错 p-重线性型, 故 (§23, 第 5 小节, 公式 (29)) 有

$$
\det(\mathrm{X}) = \mathrm{D}(v(e_1), \cdots, v(e_p), b_{p+1}, \cdots, b_{p+q}) = \det(v) \cdot \mathrm{D}(e_1, \cdots, e_p, b_{p+1}, \cdots, b_{p+q}),
$$

这样就证明了

$$\det(X) = \det(A) \cdot D(e_1, \cdots, e_p, b_{p+1}, \cdots, b_{p+q}).$$

而

$$u(e_{p+1}) = w(e_{p+1}) + a_{p+1}, \cdots, u(e_{p+q}) = w(e_{p+q}) + a_{p+q},$$

其中的向量 $a_{p+j} \in E(1 \leqslant j \leqslant q)$; 故有

$$\begin{aligned} D(e_1, \cdots, b_{p+q}) &= D(e_1, \cdots, e_p, w(e_{p+1}) + a_{p+1}, \cdots, w(e_{p+q}) + a_{p+q}) \\ &= D(e_1, \cdots, e_p, w(e_{p+1}), \cdots, w(e_{p+q})), \end{aligned}$$

这是因为 $a_{p+j} \in E$ 是 $e_1, \cdots, e_p$ 的线性组合 (比如说利用第 1 小节的 e)). 但与前面的推理类似的推理显然指出

$$D(e_1, \cdots, e_p, w(e_{p+1}), \cdots, w(e_{p+q})) = \det(w) \cdot D(e_1, \cdots, e_{p+q}),$$

而由于 $\det(w) = \det(D)$, 终于得到我们要证明的公式

$$\det(X) = \det(A) \det(D).$$

3. 伴随矩阵

设

$$X = \begin{pmatrix} \xi_{11} & \cdots & \xi_{1n} \\ \vdots & & \vdots \\ \xi_{n1} & \cdots & \xi_{nn} \end{pmatrix}$$

是元素在一个交换环 K 内的方阵. 称矩阵

$$\widetilde{X} = \begin{pmatrix} \tilde{\xi}_{11} & \cdots & \tilde{\xi}_{1n} \\ \vdots & & \vdots \\ \tilde{\xi}_{n1} & \cdots & \tilde{\xi}_{nn} \end{pmatrix}$$

为 X 的伴随矩阵, $\widetilde{X}$ 的元素由关系

$$\tilde{\xi}_{ij} = (-1)^{i+j} \det(X_{ji}) \tag{9}$$

给定, 我们再提醒 X_{ij} 是从 X 划去第 i 行和第 j 列得到的矩阵.

定理 2 设 X 是元素在一个交换环内的一个 n 阶方阵, 则有

$$\widetilde{X} \cdot X = X \cdot \widetilde{X} = \det(X) \cdot I_n.$$

例如证明 $\widetilde{X}\cdot X=\det(X)\cdot I_n$, 即 $X\cdot\widetilde{X}$ 是对角矩阵, 并且对角线上的元素都等于 $\det(X)$, 这归结为证明

$$\sum_j \tilde{\xi}_{ij}\xi_{jk}=\begin{cases}\det(X), & i=k,\\ 0, & i\neq k.\end{cases}\tag{10}$$

而左端写成

$$\sum_j (-1)^{i+j}\xi_{jk}\det(X_{ji}).$$

根据公式 (7), 这个表达式刚好是 X 的第 i 列的元素 $\xi_{1i},\cdots,\xi_{ni}$ 用 $\xi_{1k},\cdots,\xi_{nk}$ 代换所得到的矩阵的行列式. 如果 $i\neq k$, 那么这样得到的矩阵有两行 (i 行和 k 行) 相等, 故其行列式是零; 如果 $i=k$, 所得到的矩阵正是 X. 由此即得 (10). 公式 $X\cdot\widetilde{X}=\det(X)\cdot I_n$ 的证明以类似方式进行: 这时必须利用 (8) 代替 (7)

推论 1 设 X 是元素在一个交换环 K 内的方阵. X 是可逆的, 必须并且只需 X 的行列式 $\det(X)$ 在 K 内是一个可逆元素; 这时有

$$X^{-1}=\det(X)^{-1}\cdot\widetilde{X}.\tag{11}$$

如果 X 是可逆的, §23 的定理 7 表明

$$\det(X\cdot X^{-1})=\det(X)\det(X^{-1})=\det(I_n)=1,$$

故 $\det(X)$ 在 K 内是可逆的. 反之, 设这个条件满足, 定理 2 显然表明矩阵

$$\det(X)^{-1}\cdot\widetilde{X}$$

是 X 的逆矩阵.



注 1 交换域 K 的情形已经在 §23 处理过了 (定理 8 的推论 1).

推论 2 设 X 是一个交换环 K 上的有限生成自由模, 而 $(a_1,\cdots,a_n)$ 是 X 的一个基. 向量 $x_1,\cdots,x_n\in X$ 组成 X 的一个基, 必须并且只需它们关于基 $(a_1,\cdots,a_n)$ 的行列式是 K 中的可逆元素.

事实上, 向量 $x_1,\cdots,x_n\in X$ 组成 X 的一个基就表示 x_i 关于所说的基的坐标组成一个在环 $M_n(K)$ 内的一个可逆矩阵.

例 2 取 $K=\mathbf{Z}$, 而 $X=\mathbf{Z}^n$, 基 (a_i) 是典范基. 我们得到下列结果: 向量

$$x_j=(\xi_{1j},\cdots,\xi_{nj})\quad(1\leqslant j\leqslant n)$$

组成 $\mathbf{Z}^n$ 的一个基, 必须并且只需

$$\det((\xi_{ij}))=\pm 1.$$

我们同样发现群 $GL(n,\mathbf{Z})$ 由元素在 $\mathbf{Z}$ 内的行列式为 1 或 -1 的矩阵组成.

4. Cramer 公式

我们已经看到系数在一个域 K 内的线性方程组的求解总可以归结为一个 Cramer 方程组的求解, 这个方程组就是

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{1n}\xi_n = \beta_1, \\ \qquad\cdots\cdots\cdots\cdots \\ \alpha_{n1}\xi_1 + \cdots + \alpha_{nn}\xi_n = \beta_n, \end{cases} \tag{12}$$

其中的矩阵

$$\mathrm{A} = (\alpha_{ij})_{1\leqslant i,j\leqslant n}$$

是可逆的, 即其行列式非零. 令

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \quad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix},$$

给定的方程组可以写成

$$\mathrm{A}x = b,$$

并且有解

$$x = \mathrm{A}^{-1}b.$$

而我们在前一个小节 (定理 2 的推论 1) 已经给出了一个计算 A 的逆矩阵的明晰公式. 如果使用这个公式, 我们便得到解

$$x = \det(\mathrm{A})^{-1} \cdot \widetilde{\mathrm{A}}b.$$

进一步把此式明晰化即得

$$\det(\mathrm{A}) \cdot \xi_i = \sum_j (-1)^{i+j} \det(\mathrm{A}_{ji})\beta_j,$$

其中 A_{ji} 是由 A 去掉第 j 行和第 i 列得到的矩阵. 而根据定理 1, 这个关系的右端正是一个这样的矩阵的行列式, 它从 A 出发, 把第 i 列的元素 $\alpha_{i1}, \cdots, \alpha_{in}$ 换成方程组 (12) 的右端 $\beta_1, \cdots, \beta_n$ 而得到. 因此 Cramer 方程组 (12) 的解由以下公式给定:

$$\xi_i = \frac{\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1,i-1} & \beta_1 & \alpha_{1,i+1} & \cdots & \alpha_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{n,i-1} & \beta_n & \alpha_{n,i+1} & \cdots & \alpha_{nn} \end{vmatrix}}{\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{vmatrix}}.$$

这就是著名的 **Cramer 公式**.

例 3 取 K $=\mathbf{R}$, 方程组是

$$\begin{cases} 2x+3y+4z=a, \\ 5x+6y+7z=b, \\ 8x+9y+9z=c, \end{cases}$$

方程组的行列式是

$$\begin{vmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 9 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 1 \\ 5 & 1 & 1 \\ 8 & 1 & 0 \end{vmatrix} = \begin{vmatrix} -3 & 0 & 0 \\ -3 & 0 & 1 \\ 8 & 1 & 0 \end{vmatrix} = 3,$$

故方程组是 Cramer 方程组. 未知元 x 由

$$x = \frac{1}{3}\begin{vmatrix} a & 3 & 4 \\ b & 6 & 7 \\ c & 9 & 9 \end{vmatrix} = \frac{1}{3}\begin{vmatrix} a & 3 & 1 \\ b & 6 & 1 \\ c & 9 & 0 \end{vmatrix} = -3a+3b-c$$

给定.



注 2 如果 K 是一个交换环, Cramer 公式必然还是有效的, 只要方程组的行列式在 K 内是可逆的.

§24 习题

计算下列行列式:

1. $\begin{vmatrix} a & 3 & 0 & 5 \\ 0 & b & 0 & 2 \\ 1 & 2 & c & 3 \\ 0 & 0 & 0 & d \end{vmatrix}$.

2. $\begin{vmatrix} x & a & b & 0 & c \\ 0 & y & 0 & 0 & d \\ 0 & e & z & 0 & f \\ g & h & k & u & l \\ 0 & 0 & 0 & 0 & v \end{vmatrix}$.

3. $\begin{vmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{vmatrix}$.

4. $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{vmatrix}$.

5. $\begin{vmatrix} 6 & -5 & 8 & 4 \\ 9 & 7 & 5 & 2 \\ 7 & 5 & 3 & 7 \\ -4 & 8 & -8 & -3 \end{vmatrix}$.

6. $\begin{vmatrix} 24 & 11 & 13 & 17 & 19 \\ 51 & 13 & 32 & 40 & 46 \\ 61 & 11 & 14 & 50 & 56 \\ 62 & 20 & 7 & 13 & 52 \\ 80 & 24 & 45 & 57 & 70 \end{vmatrix}$.

7. $\begin{vmatrix} 1 & 2 & 3 & \cdots & n \\ -1 & 0 & 3 & \cdots & n \\ -1 & -2 & 0 & \cdots & n \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & -2 & -3 & \cdots & 0 \end{vmatrix}$. 8. $\begin{vmatrix} a_0 & a_1 & a_2 & \cdots & a_n \\ -x & x & 0 & \cdots & 0 \\ 0 & -x & x & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & -x & \cdots & x \end{vmatrix}$.

¶9. $\begin{vmatrix} 3 & 2 & 0 & 0 & \cdots & 0 \\ 1 & 3 & 2 & 0 & \cdots & 0 \\ 0 & 1 & 3 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 3 \end{vmatrix}$.

[用 D_n 表示这个 n 阶行列式, 建立 $\mathrm{D}_n, \mathrm{D}_{n-1}$ 和 D_{n-2} 之间的简单关系, 并用下面这个关系: 设 $(u_n)_{n\geqslant 1}$ 是一个复数序列, 使得有递推关系

$$u_{n+2} = au_{n+1} + bu_n;$$

设 z_1 和 z_2 是方程

$$z^2 - az - b = 0$$

的 (相等或不等) 解. 如果 $z_1 \neq z_2$, 则存在常数 c_1 和 c_2, 使得

$$u_n = c_1 z_1^n + c_2 z_2^n \quad \text{对于所有 } n,$$

如果 $z_1 = z_2$, 则存在常数 c_1 和 c_2, 使得

$$u_n = (c_1 n + c_{)} z_1^n \quad \text{对于所有 } n.$$

对于更一般的结果, 参见 §35, 习题 16.]

¶10. $\begin{vmatrix} 1 & 2 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 3 & 4 & 3 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 2 & 5 & 3 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & 5 & 3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 5 & 3 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 2 & 5 \end{vmatrix}$ (与上题同样的方法).

11. $\begin{vmatrix} 1-n & 1 & 1 & \cdots & 1 \\ 1 & 1-n & 1 & \cdots & 1 \\ 1 & 1 & 1-n & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1-n \end{vmatrix}$ (n 行 n 列的行列式).

12. $\begin{vmatrix} 1 & n & n & \cdots & n \\ n & 2 & n & \cdots & n \\ n & n & 3 & \cdots & n \\ \vdots & \vdots & \vdots & & \vdots \\ n & n & n & \cdots & n \end{vmatrix}$.

13. $\begin{vmatrix} x & a_1 & a_2 & \cdots & a_{n-1} & 1 \\ a_1 & x & a_2 & \cdots & a_{n-1} & 1 \\ a_1 & a_2 & x & \cdots & a_{n-1} & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_1 & a_2 & a_3 & \cdots & x & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n & 1 \end{vmatrix}$ (求 x 的这个多项式函数的根).

¶14. 证明

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1\leqslant i<j\leqslant n} (x_j - x_i).$$

(Vandermond 行列式; 注意, 例如, 左端是 x_1 的至多 $n-1$ 次的多项式, $x_2, \cdots, x_n$ 显然是它的根).

¶15. $\begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-2} & x_1^n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-2} & x_n^n \end{vmatrix}$.

¶16. $\begin{vmatrix} 1 & f_1(x_1) & f_2(x_1) & \cdots & f_{n-1}(x_1) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & f_1(x_n) & f_2(x_n) & \cdots & f_{n-1}(x_n) \end{vmatrix}$, 其中 $f_k(x) = x^k + a_{k1}x^{k-1} + \cdots + a_{kk} (1 \leqslant k \leqslant n-1)$.

¶17. $\begin{vmatrix} 1 & \binom{x_1}{1} & \binom{x_1}{2} & \cdots & \binom{x_1}{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \binom{x_n}{1} & \binom{x_n}{2} & \cdots & \binom{x_n}{n-1} \end{vmatrix}$, 其中 $\binom{x}{k} = \dfrac{x(x-1)\cdots(x-k+1)}{k!}$.

¶¶18. 证明

$$
\begin{vmatrix}
a & b & c & d & e & f & g & h \\
b & a & d & c & f & e & h & g \\
c & d & a & b & g & h & e & f \\
d & c & b & a & h & g & f & e \\
e & f & g & h & a & b & c & d \\
f & e & h & g & b & a & d & c \\
g & h & e & f & c & d & a & b \\
h & g & f & e & d & c & b & a
\end{vmatrix}
$$

$$
\begin{aligned}
&= (a+b+c+d+e+f+g+h)(a+b+c+d-e-f-g-h) \\
&\quad \cdot(a+b-c-d+e+f-g-h)(a+b-c-d-e-f+g+h) \\
&\quad \cdot(a-b+c-d+e-f+g-h)(a-b+c-d-e+f-g+h) \\
&\quad \cdot(a-b-c+d+e-f-g+h)(a-b-c+d-e+f+g-h).
\end{aligned}
$$

19. 证明如果 $n \geqslant 3$, 则

$$
\begin{vmatrix}
1+x_1y_1 & 1+x_1y_2 & \cdots & 1+x_1y_n \\
\vdots & \vdots & & \vdots \\
1+x_ny_1 & 1+x_ny_2 & \cdots & 1+x_ny_n
\end{vmatrix} = 0.
$$

¶20. 用关于行数 n 的归纳法计算行列式

$$
\begin{vmatrix}
1 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 1
\end{vmatrix}.
$$

¶21. 证明

$$
\begin{vmatrix}
1 & 1 & 0 & 0 & \cdots & 0 \\
1 & \binom{2}{1} & \binom{2}{2} & 0 & \cdots & 0 \\
1 & \binom{3}{1} & \binom{3}{2} & \binom{3}{3} & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & & \vdots \\
1 & \binom{n}{1} & \binom{n}{2} & \binom{n}{3} & \cdots & \binom{n}{n-1}
\end{vmatrix} = 1.
$$

¶22. 通过计算乘积

$$
\begin{vmatrix}
a & -b & -c & -d \\
b & a & -d & c \\
c & d & a & -b \\
d & -c & b & a
\end{vmatrix}
\begin{vmatrix}
x & -y & -z & -t \\
y & x & -t & z \\
z & t & x & -y \\
t & -z & y & x
\end{vmatrix}
$$

建立 Euler 等式

$$\begin{aligned}&(a^2+b^2+c^2+d^2)(x^2+y^2+z^2+t^2)\\=&(ax+by+cz+dt)^2+(ay-bx+ct-dz)^2\\&+(az-bt-cx+dy)^2+(at+bz-cy+dx)^2.\end{aligned}$$

你发现同 §15 的习题 10 和 11 的关系了吗?

计算下列矩阵的逆矩阵:

23. $\begin{pmatrix}2&7&3\\3&9&4\\1&5&3\end{pmatrix}$.

24. $\begin{pmatrix}1&2&3&4\\2&3&1&2\\1&1&1&-1\\1&0&-2&-6\end{pmatrix}$.

25. $\begin{pmatrix}3&-2&-5&1\\1&-3&1&5\\1&2&0&-4\\1&-1&-4&9\end{pmatrix}$.

26. 计算矩阵的秩

$$\begin{pmatrix}17&-28&45&11&39\\24&-37&61&13&50\\25&-7&32&-18&-11\\31&12&19&-43&-55\\42&13&29&-55&-68\end{pmatrix},\quad\begin{pmatrix}2&-1&1&3&4\\2&-1&2&1&-2\\2&-3&1&2&-2\\1&0&1&-2&-6\\1&2&1&-1&0\\4&-1&3&-1&-8\end{pmatrix}.$$

27. 下列向量

$$(1,0,0,2,5),\quad(0,1,0,3,4),\quad(0,0,1,4,7),\quad(2,-3,4,11,12)$$

在 $\mathbf{R}^5$ 线性无关吗?

¶28. 证明

$$\begin{vmatrix}2&-1&0&0&0&0&0&0\\-1&2&-1&0&0&0&0&0\\0&-1&2&-1&0&0&0&0\\0&0&-1&2&-1&0&0&0\\0&0&0&-1&2&-1&0&-1\\0&0&0&0&-1&2&-1&0\\0&0&0&0&0&-1&2&0\\0&0&0&0&-1&0&0&2\end{vmatrix}=1.$$

29. 计算

$$\begin{vmatrix} 1 & 4 & 9 & 16 \\ 4 & 9 & 16 & 25 \\ 9 & 16 & 25 & 36 \\ 16 & 25 & 36 & 49 \end{vmatrix}.$$

30. 为了计算一个十阶行列式原则上应当执行多少次加法和乘法?

31. 利用行列式理论解 §20 的习题 1 至 17.

32. 求解并且讨论下列线性方程组:

a) $\begin{cases} ax + by + z = 1, \\ x + aby + z = b, \\ x + by + az = 1. \end{cases}$

b) $\begin{cases} ax + by + 2z = 1, \\ ax + (2b-1)y + 3z = 1, \\ ax + by + (b+3)z = 2b-1. \end{cases}$

c) $\begin{cases} 2(a+1)x + 3y + az = a+4, \\ (4a-1)x + (a+1)y + (2a-1)z = 2a+2, \\ (5a-4)x + (a+1)y + (3a-4)z = a-1. \end{cases}$

¶¶33. (根据 Laplace 规则展开行列式) 设 K 是一个交换环, 而 n 是一个 >1 的整数. 用

$$\mathrm{D}(x_1, \cdots, x_n)$$

表示 n 个向量 $x_i \in \mathrm{K}^n$ 关于 K^n 的典范基 $e_1, \cdots, e_n$ 的行列式. 令

$$x_j = e_1\xi_{1j} + \cdots + e_n\xi_{nj} \quad (1 \leqslant j \leqslant n),$$

选择一个满足 $1 \leqslant p \leqslant n$ 的整数 p.

a) 证明有关系

$$\mathrm{D}(x_1, \cdots, x_n) = \sum_{i_1 < \cdots < i_p} \begin{vmatrix} \xi_{i_1 1} & \cdots & \xi_{i_1 p} \\ \vdots & & \vdots \\ \xi_{i_p 1} & \cdots & \xi_{i_p p} \end{vmatrix} \cdot \mathrm{D}(e_{i_1}, \cdots, e_{i_p}, x_{p+1}, \cdots, x_n)$$

($\mathrm{D}(x_1, \cdots, x_n)$ 看作 $x_1, \cdots, x_n$ 的交错多重线性函数).

b) 证明对于 $1 \leqslant i_1 < \cdots < i_p \leqslant n$, 有

$$\mathrm{D}(e_{i_1}, \cdots, e_{i_p}, x_{p+1}, \cdots, x_n) = \begin{vmatrix} \xi_{j_1, p+1} & \cdots & \xi_{j_1, n} \\ \vdots & & \vdots \\ \xi_{j_{n-p}, p+1} & \cdots & \xi_{j_{n-p}, n} \end{vmatrix},$$

其中的 $j_1, \cdots, j_{n-p}$ 表示 $\{1, 2, \cdots, n\}$ 中的这样的整数, 它们不属于集合 $\{i_1, \cdots, i_p\}$, 排序后使得置换 $\begin{pmatrix} 1 & \cdots & p & p+1 & \cdots & n \\ i_1 & \cdots & i_p & j_1 & \cdots & j_{n-p} \end{pmatrix}$ 是偶置换.

c) 设

$$X = (\xi_{ij})_{1\leqslant i,j\leqslant n}$$

是元素在 K 内的一个 n 阶方阵. 对于 $\{1,2,\cdots,n\}$ 的含有 p 个元素的子集 I, 用 X_I 表示这样的 ξ_{ij} 组成的矩阵, $i\in I$, 并且 $1\leqslant j\leqslant p$, 用 X_I' 表示由这样的 ξ_{ij} 组成的 "补" 矩阵, $i\notin I$, 并且 $p+1\leqslant j\leqslant n$. 最后用 $n(I)$ 表示这样的序偶 (i,j) 的数目, $i\in I$, $j\notin I$, 并且 $i>j$. 证明

$$\det(X) = \sum_{\mathrm{Card}(I)=p} (-1)^{n(I)} \det(X_I)\det(X_I')$$

(Laplace 公式), 其中的和取自 $\{1,2,\cdots,n\}$ 的含有 p 个元素的所有子集 I.

d) 由外积的结合性公式 [§23, 习题 12, d)] 推出这个结果.

利用 Laplace 规则计算下列行列式:

34. $\begin{vmatrix} 1 & 1 & 3 & 4 \\ 2 & 0 & 0 & 8 \\ 3 & 0 & 0 & 2 \\ 4 & 4 & 7 & 5 \end{vmatrix}$.

35. $\begin{vmatrix} 2 & 1 & 4 & 3 & 5 \\ 3 & 4 & 0 & 5 & 0 \\ 3 & 4 & 5 & 2 & 1 \\ 1 & 5 & 2 & 4 & 3 \\ 4 & 6 & 0 & 7 & 0 \end{vmatrix}$.

36. $\begin{vmatrix} 1 & 2 & 3 & 4 & 5 & 3 \\ 6 & 5 & 7 & 8 & 4 & 2 \\ 9 & 8 & 6 & 7 & 0 & 0 \\ 3 & 2 & 4 & 5 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 \\ 5 & 6 & 0 & 0 & 0 & 0 \end{vmatrix}$.

37. $\begin{vmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & a & b & c & d \\ 0 & a^2 & b^2 & c^2 & d^2 \end{vmatrix}$.

38. $\begin{vmatrix} 3 & 4 & -3 & -1 & 2 \\ -5 & 6 & 5 & 2 & 3 \\ 4 & -9 & -3 & 7 & -5 \\ -1 & -4 & 1 & 1 & -2 \\ -3 & 7 & 5 & 2 & 3 \end{vmatrix}$.

§25 仿射空间

1. 平移向量空间

在这一小节我们使用“空间”、“点”、“向量”、“等价”、“平移”等词汇, 读者应当赋予它们初等几何中的同样的意义. 鉴于像空间这样的概念, 严格来说并不是数学对象 (我们不能借助基本符号和集合论描述它), 我们不给以精确的定义. 不过人们可以以某种方式通过经验获得它们的性质, 这些性质可以作为构建类似的数学对象即仿射空间的基础, 仿射空间将在下一小节定义.

用 E 表示通常的三维空间, 其元素是通常的点. 我们曾经看到 (§10, 例 2) 如果一次性地选定 E 的一个点 O, 可以把 E 内的起点在 O 的向量的集合看作一个三维实向量空间. 但是这个向量空间的定义包含一个任意元素, 即点 O 的选择, 如果 O 用另一个点 O′ 代替, 所得到的向量空间同构于但并不全等于第一个 [令起点为 O 的所有向量对应于起点为 O′ 的与之相等 (包含平行) 的向量, 就得到从第一个空间到第二个的典范同构]. 我们现在指出通过修改这个构建, 可以用一个通常的空间对应一个典范向量空间, 这个空间的定义不再包含任何任意的选择.

为此用 T 表示 E 内的所有平移的集合: 一个平移是一个从 E 到 E 内的双射, 它使每一个点 $\mathrm{P} \in \mathrm{E}$ 变换到点 $\mathrm{P}' \in \mathrm{E}$, 使得向量 PP′ 等于一个给定的向量. 我们要指出集合 T 可以配备一个实向量空间结构.

为此应当定义两个平移的和 $s+t$ 以及一个平移与一个标量 $\lambda \in \mathbf{R}$ 的乘积 λs. 关于和, 我们令

$$s+t=s\circ t,$$

这平移 s 和 t 的复合. 至于 λs, 这将是一个这样的平移, 其向量由 s 的向量乘以 λ 而得到.

选择一个点 $\mathrm{O} \in \mathrm{E}$, 令所有起点为 O 的向量 x 对应向量为 x 的平移 s_x: 对于所有的点 $\mathrm{P} \in \mathrm{E}$, 起点为 P、终点为 $s_x(\mathrm{P})$ 的向量等于 x. 从这些定义看出显然有

$$s_x+s_y=s_{x+y}, \quad \lambda\cdot s_x=s_{\lambda x}. \tag{1}$$

由于起点为 O 的向量集合配备显然的代数运算之后是一个实向量空间, 那么集合 T 配备了前面所定义的运算也是一个向量空间, 并且事实上映射 $x\to s_x$ 是从第一个空间到第二个空间上的一个同构.

我们说 T 是在通常空间内的平移向量空间.

给定一个点 $\mathrm{P} \in \mathrm{E}$ 和一个平移 $s \in \mathrm{T}$, 作为定义我们令

$$s+\mathrm{P}=s(\mathrm{P}),$$

这样我们就得到了一个从 $\mathrm{T}\times\mathrm{E}$ 到 E 内的一个映射 $(s,\mathrm{P})\to s+\mathrm{P}$, 并且显然有下列性质:

(EA1) 对于任意 $s, t \in \mathrm{T}$ 和 $\mathrm{P} \in \mathrm{E}$ 有

$$s + (t + \mathrm{P}) = (s + t) + \mathrm{P}.$$

(EA2) 对于所有 $\mathrm{P} \in \mathrm{E}$ 有

$$0 + \mathrm{P} = \mathrm{P}.$$

(在这个公式里, 0 必然是 T 的中性元.)

(EA3) 对于任意 $\mathrm{P}, \mathrm{Q} \in \mathrm{E}$, 存在唯一的一个 $s \in \mathrm{T}$, 使得

$$s + \mathrm{P} = \mathrm{Q}.$$

当我们有

$$s + \mathrm{P} = \mathrm{Q}$$

时, 经常作为定义写出

$$s = \mathrm{Q} - \mathrm{P};$$

两个点 P 和 Q 之间的 "差" 是把 P 拉到 Q 的平移.



注 1 请注意, 我们刚定义了两个点的 "差", 它是一个平移而非一个点或向量, 但反之鉴于没有任何物理的和几何的意义我们没有定义两个点的 "和".

2. 与一个向量空间相伴的仿射空间

设 T 是一个域 K 上的向量空间. 称由一个集合 E 和一个记作

$$(s, \mathrm{P}) \to s + \mathrm{P}$$

的从 $\mathrm{T} \times \mathrm{E}$ 到 E 内的映射组成的对为**与 T 相伴的仿射空间**, 如果前一小节的性质 (EA1), (EA2) 和 (EA3) 满足. 这时称 E 的元素为**点**, 而所有 $s \in \mathrm{T}$ 允许定义一个从 E 到 E 内的映射, 即一个我们记作 $\bar{s}$ 的映射, 其定义是

$$\bar{s}(\mathrm{P}) = s + \mathrm{P} \quad 对于所有\ \mathrm{P} \in \mathrm{E}.$$

这个映射 $\bar{s}$ 是双射, 把它叫作 E 内的**平移**.

例 1 设 T 是域 K 上的一个向量空间. 我们可以把 T 本身看作与 T 相伴的一个仿射空间, 鉴于 T 内的加法是从 $\mathrm{T} \times \mathrm{T}$ 到 T 内的一个映射, 它显然满足条件 (EA1), (EA2) 和 (EA3) —— 这些条件在这种情形, 只是等于说 T 配备了加法是一个群.

例 2 设 X 是一个集合, A 是 X 的一个子集, 而 u 是从 A 到域 K 内的一个给定的映射. 用 T 表示从 X 到 K 内的这样的映射 s 的集合: 对于所有 $a \in \mathrm{A}$ 有 $s(a)=0$; 而 E 是从 X 到 K 内的这样的映射 f 的集合: 对于所有 $a \in \mathrm{A}$ 有 $f(a)=u(a)$ (即 E 是 u 到 X 的延拓的集合). 对于 $s \in \mathrm{T}$ 和 $f \in \mathrm{E}$ 定义 $s+f$ 为函数 $s(x)+f(x)$; 第 1 小节的条件是满足的, 所以可以把 E 看作与 T 相伴的仿射空间.

读者将会证实这个例子是下一个例子的特殊情形.

例 3 设 M 是 K 上的一个向量空间, 取 M 的一个向量子空间作为 T. 由于 T 是加法群 M 的一个子群, 故可以考虑在 M 内 T 的陪集 (§7, 第 6 小节). 对于所有 $a \in \mathrm{M}$, 用 $\mathrm{T}+a$ 表示 T 包含 a 的陪集: 这是 M 的一个子集, 即使得 $x-a \in \mathrm{T}$ 的 $x \in \mathrm{M}$ 的集合.

设 E 是 M 内 T 的陪集, 一般这不是 M 的向量子空间, 但是对于 $s \in \mathrm{T}$ 和 $u \in \mathrm{E}$, 显然和向量 $s+u$ 仍然在 E 内. 由此得到从 $\mathrm{T} \times \mathrm{E}$ 到 E 内的一个映射, 而这个映射使得我们可以把 E 看作相伴于 T 的一个仿射空间, 这个事实立刻可以验证.

例 4 设 E 和 T 跟第 1 小节中的一样 (即 E 是通常的空间和 T 是 E 内平移向量空间). 设 $\mathrm{E}' \subset \mathrm{E}$ 是一张平面 (对应的, 一条直线), 设 $\mathrm{T}' \subset \mathrm{T}$ 是平行于 E' 的向量的平移, 即映射 E' 到 E' 内的平移的集合. 显然 T' 是 T 的一个向量子空间, 并且对于 $s \in \mathrm{T}'$ 和 $\mathrm{P} \in \mathrm{E}'$ 有 $s+\mathrm{P} \in \mathrm{E}'$, 可以典范地定义从 $\mathrm{T}' \times \mathrm{E}'$ 到 E' 内的一个映射. 像容易看到的那样, 这就让我们可以把 E' 看作相伴于 T' 的一个仿射空间.

设 T 是 K 上的一个向量空间, 而 E 是相伴于 T 的一个仿射空间. 选择 E 的一个点 O, 我们要指出可以把 E 看作 K 上的一个向量空间 (不过 K 上的这个向量空间的构造依赖点 O 的选择, 即不是典范的). 为此, 给定 E 的两个点 P 和 Q, 令 $\mathrm{P}+\mathrm{Q}=\mathrm{R}$, 其中 R 使得下式成立:

$$\mathrm{R}-\mathrm{O}=(\mathrm{P}-\mathrm{O})+(\mathrm{Q}-\mathrm{O}) \tag{2}$$

(这意味着把 O 拉到 R 的平移由把 O 拉到 P 的平移和把 O 拉到 Q 的平移合成). 给定 E 的一个点 P 和 K 的一个标量 λ, 用关系

$$\mathrm{P}'-\mathrm{O}=\lambda \cdot (\mathrm{P}-\mathrm{O}) \tag{3}$$

定义点 $\mathrm{P}'=\lambda\mathrm{P}$. 我们要指出, 配备了这些运算, E 是一个同构于 T 的向量空间.

为此, 考虑由

$$f(s)=s+\mathrm{O}$$

定义的映射 $f: \mathrm{T} \to \mathrm{E}$; 根据条件 (EA3), 它是双射. 设 $s, t \in \mathrm{T}$, 令 $\mathrm{P}=f(s)$ 和 $\mathrm{Q}=f(t)$, 我们来计算由 (2) 给定的点 $\mathrm{R}=\mathrm{P}+\mathrm{Q}$; 我们有 $\mathrm{P}=s+\mathrm{O}$, 故 $\mathrm{P}-\mathrm{O}=s$, 同

样有 $Q - O = t$, 于是根据 (2) 有 $R - O = s + t$, 而这就是说 $R = f(s + t)$, 故得

$$f(s + t) = f(s) + f(t).$$

我们发现借助 (3) 可以得到

$$f(\lambda s) = \lambda f(s).$$

由于向量空间的公理在 T 内是满足的, 又由于 f 是双射, 故向量空间的公理在 E 内也满足, 并且 f 是向量空间的一个同构.

给集合 E 配备了运算 (2) 和 (3) 所得到的向量空间将记作 E_O.

例 5 取第 1 小节中的 E 和 T. 选定一个 "原点" O, E 内的加法 $P + Q = R$ 由向量 $\overrightarrow{OR}$ 是向量 $\overrightarrow{OP}$ 和 $\overrightarrow{OQ}$ 的和这个条件定义, 而点 $\lambda P = P'$ 由向量 $\overrightarrow{OP'}$ 等于向量 $\overrightarrow{OP}$ 乘以 λ (即 P' 是在中心为 O 比例为 λ 的位似变换下的像) 定义.

3. 仿射空间内的重心

假设和记号保持与第 2 小节同样, 选择一个 "起点" 后在 E 上定义一个向量空间, 我们要考查这个向量空间的结构对于起点的依赖.

为此考虑点 $P_1, \cdots, P_n \in E$ 和标量 $\lambda_1, \cdots, \lambda_n \in K$. 一旦选定一个点 $O \in E$, 我们可以在向量空间 E_O 内定义线性组合

$$P = \lambda_1 P_1 + \cdots + \lambda_n P_n.$$

考虑到关系 (2) 和 (3) 显然有

$$P - O = \lambda_1(P_1 - O) + \cdots + \lambda_n(P_n - O), \tag{4}$$

这个关系是在向量空间 T 的元素之间的一个线性关系.

如果用另一个点 O' 代替 O, 点 P 则用由

$$P' - O' = \lambda_1(P_1 - O') + \cdots + \lambda_n(P_n - O') \tag{5}$$

定义的 P' 代替. 而给定了 E 的点 A, B, C, 则以一般的方式有关系

$$A - C = (A - B) + (B - C), \tag{6}$$

因为如果令 $s = A - B$ 和 $t = B - C$, 则有 $A = s + B$ 和 $B = t + C$, 故 $A = s + (t + C)$, 根据 (EA1), 因此有 $A = (s + t) + C$, 于是正如所宣布的 $A - C = s + t$. 据此我们有

$$P_i - O' = (P_i - O) - (O' - O).$$

由 (5) 得到

$$\begin{aligned}
P' - O' &= \lambda_1[(P_1 - O)+(O - O') + \cdots + \lambda_n[(P_n - O)+(O - O')] \\
&= \lambda_1(P_1 - O) + \cdots + \lambda_n(P_n - O) + (\lambda_1 + \cdots + \lambda_n) \cdot (O - O') \\
&= P - O + (\lambda_1 + \cdots + \lambda_n)(O - O'),
\end{aligned}$$

由于 $P' - O' = (P' - P) + (P - O)$, 这个关系可以改写为

$$P' - P = (\lambda_1 + \cdots + \lambda_n - 1)(O - O'). \tag{7}$$

这个关系表明一般说来 $P \neq P'$, 即 E 的向量空间结构实际上依赖原点 O 的选择.

不过注意到一旦

$$\lambda_1 + \cdots + \lambda_n = 1, \tag{8}$$

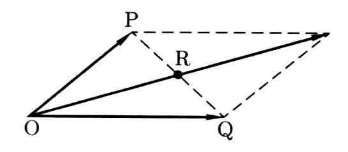
就有 $P = P'$, 即当条件 (8) 满足时, 点 $\lambda_1 P_1 + \cdots + \lambda_n P_n$ 首先要在 E 内选取一个原点 O 才有定义, 而事实上又不依赖 O 的选取, 由此具有 "内蕴的" 或 "典范的" 意义. 称这个点为分配了质量 $\lambda_1, \cdots, \lambda_n$ 的点 $P_1, \cdots, P_n$ 的**质心** (或**重心**).

例 6 取第 1 小节中的 E 和 T. 给定两个点 P 和 Q, 可以定义点

$$R = \frac{1}{2}P + \frac{1}{2}Q,$$

在实际中写成

$$\frac{P + Q}{2}.$$



它由关系

$$\overrightarrow{OR} = \frac{1}{2}(\overrightarrow{OP} + \overrightarrow{OQ})$$

给定, 其中 O 是 E 内任意一个点. 由于 $\overrightarrow{OP} + \overrightarrow{OQ}$ 由平行四边形法则给定, 故显然 R 是连接 P 和 Q 的直线段的中点.

更一般的, 考虑点

$$R = tP + (1 - t)Q,$$

其中 t 是任意一个实数, R 由关系

$$\overrightarrow{OR} = t \cdot \overrightarrow{OP} + (1 - t) \cdot \overrightarrow{OQ} = \overrightarrow{OQ} + t(\overrightarrow{OP} - \overrightarrow{OQ})$$

给定. 现在 $\overrightarrow{\mathrm{OP}}-\overrightarrow{\mathrm{OQ}}$ 就是从原点出发, 与 $\overrightarrow{\mathrm{PQ}}$ 相等 (包含平行) 的向量. 从上面的图直接推出 R 位于过 P 和 Q 的直线上, 这条直线正是对于所有 $t\in\mathbf{R}$ 点 $t\mathrm{P}+(1-t)\mathrm{Q}$ 的集合.

例 7　仍然取 E 和 T 跟前面的一样, 再取三个点 P, Q, R. 可以定义

$$\mathrm{M}=u\mathrm{P}+v\mathrm{Q}+w\mathrm{R},\ 其中\ u+v+w=1. \tag{9}$$

如果 O 是 E 的任意一个点, 我们有

$$\begin{aligned}\overrightarrow{\mathrm{OM}}&=u\cdot\overrightarrow{\mathrm{OP}}+v\cdot\overrightarrow{\mathrm{OQ}}+w\cdot\overrightarrow{\mathrm{OR}}=u\cdot\overrightarrow{\mathrm{OP}}+v\cdot\overrightarrow{\mathrm{OQ}}+(1-u-v)\cdot\overrightarrow{\mathrm{OR}}\\&=\overrightarrow{\mathrm{OR}}+u\cdot(\overrightarrow{\mathrm{OP}}-\overrightarrow{\mathrm{OR}})+v\cdot(\overrightarrow{\mathrm{OQ}}-\overrightarrow{\mathrm{OR}}),\end{aligned}$$

而 $(\overrightarrow{\mathrm{OM}}-\overrightarrow{\mathrm{OR}})$ 等于 $\overrightarrow{\mathrm{RM}}$, $(\overrightarrow{\mathrm{OP}}-\overrightarrow{\mathrm{OR}})$ 等于 $\overrightarrow{\mathrm{RP}}$, $(\overrightarrow{\mathrm{OQ}}-\overrightarrow{\mathrm{OR}})$ 等于 $\overrightarrow{\mathrm{RQ}}$. 我们发现向量 $\overrightarrow{\mathrm{RM}}$ 是向量 $\overrightarrow{\mathrm{RP}}$ 和 $\overrightarrow{\mathrm{RQ}}$ 的线性组合, 更准确地有

$$\overrightarrow{\mathrm{RM}}=u\cdot\overrightarrow{\mathrm{RP}}+v\overrightarrow{\mathrm{RQ}},$$

因此 M 位于由三个点 P, Q, R 所决定的平面上 (假定三个点不位于一条直线上). 反之显然这个平面上的所有点对于 u 和 v 的适当选择都可以写成 (9) 这种形式. 在特殊情形, 取 $u=v=w=\dfrac{1}{3}$, 点 M 由

$$\overrightarrow{\mathrm{RM}}=\frac{1}{3}(\overrightarrow{\mathrm{RP}}+\overrightarrow{\mathrm{RQ}})=\frac{2}{3}\cdot\frac{\overrightarrow{\mathrm{RP}}+\overrightarrow{\mathrm{RQ}}}{2}$$

给定. 这正是三角形 PQR 的在通常意义下的重心.

当计算仿射空间内的质心时, 经常需要分配律

$$\sum_i\mu_i\sum_j\lambda_{ij}\cdot\mathrm{P}_{ij}=\sum_{i,j}\mu_i\lambda_{ij}\cdot\mathrm{P}_{ij}, \tag{10}$$

只要它有意义, 即当

$$\sum_j\lambda_{ij}=1,\ 对于所有\ i,\ \sum_i\mu_i=1,$$

它就成立 (注意到这些关系蕴含

$$\sum_{i,j}\mu_i\lambda_{ij}=1,$$

当上述关系成立时, 实际上 (10) 右端有意义). 为了证明 (10), 只需指出如果 O 是 E 的一个点, 则有

$$\sum_i\mu_i\left(\sum_j\lambda_{ij}\cdot\mathrm{P}_{ij}-\mathrm{O}\right)=\sum_{i,j}\mu_i\lambda_{ij}(\mathrm{P}_{ij}-\mathrm{O}),$$

而根据定义我们有

$$\sum_j \lambda_{ij} \cdot \mathrm{P}_{ij} - \mathrm{O} = \sum_j \lambda_{ij} \cdot (\mathrm{P}_{ij} - \mathrm{O}),$$

因此要建立的关系无非是向量空间 T 的分配律.

4. 仿射空间内的线性流形

初等几何“直线”和“平面”的概念可以推广如下. 设 T 是域 K 上的一个向量空间, E 是一个相伴于 T 的仿射空间, 而 V 是 E 的一个子集. 称 V 是 E 内的一个**线性流形**, 如果对于任意点 $\mathrm{P}_1, \cdots, \mathrm{P}_n \in \mathrm{V}$ 和标量 $\lambda_1, \cdots, \lambda_n \in \mathrm{K}$, 只要

$$\lambda_1 + \cdots + \lambda_n = 1$$

就有

$$\lambda_1 \mathrm{P}_1 + \cdots + \lambda_n \mathrm{P}_n \in \mathrm{V}.$$

例 8 空集是一个线性流形.

例 9 对于所有点 $\mathrm{P} \in \mathrm{E}$, 缩减为一个点 P 的集合 $\{\mathrm{P}\}$ 是一个线性流形, 这来自于这样的事实,

$$\lambda_1 \mathrm{P}_1 + \cdots + \lambda_n \mathrm{P}_n = \mathrm{P}, \text{ 如果 } \mathrm{P}_1 = \cdots = \mathrm{P}_n = \mathrm{P}.$$

在实际中, 人们对于点 P 和线性流形 $\{\mathrm{P}\}$ 自然不加以区别.

例 10 设 $\mathrm{P}_1, \cdots, \mathrm{P}_n$ 是 E 内给定的点, 那么可以通过其和为 1 的标量 $\lambda_i \in \mathrm{K}$ 表示成形式

$$\mathrm{P} = \lambda_1 \mathrm{P}_1 + \cdots + \lambda_n \mathrm{P}_n$$

的 $\mathrm{P} \in \mathrm{E}$ 的集合 V 是一个线性流形. 事实上, 给定 V 的点

$$\mathrm{Q}_j = \sum_i \lambda_{ij} \mathrm{P}_i, \text{ 其中 } \sum_i \lambda_{ij} = 1,$$

以及其和为 1 的标量 μ_i, 根据 (10) 有

$$\sum_j \mu_j \mathrm{Q}_j = \sum_j \mu_j \sum_i \lambda_{ij} \mathrm{P}_i = \sum_{i,j} \mu_i \lambda_{ij} \mathrm{P}_i = \sum_i \nu_i \mathrm{P}_i,$$

其中

$$\nu_i = \sum_j \mu_j \lambda_{ij}.$$

这就表明 V 满足线性流形的定义.

我们注意到 V 含给定的点 P_i (比如, 取 $\lambda_1 = 1$ 和 $\lambda_2 = \cdots = \lambda_n = 0$ 就得到关系 $P_1 \in V$), 并且根据定义所有含有点 P_i 的线性流形包含 V. 因此, V 是含有点 P_i 的最小线性流形. 称这个线性流形为**由点** $P_1, \cdots, P_n$ **生成的线性流形**.

例 11　设 P 和 Q 是 E 的两个点, 它们生成的线性流形由点 $\lambda P + (1-\lambda)Q$ 组成, 其中 $\lambda \in K$ 是任意的. 如果 $P = Q$, 它缩减为一个点 P. 如果 $P \neq Q$, 称它为**连接 P 和 Q 的直线**. 这个术语由例 6 说明其合理性.

我们注意, 如果 K 是模 2 整数的域, 这个域仅有元素 0 和 1, 连接 P 和 Q 的直线缩减为集合 $\{P, Q\}$. 这种状况无疑让初学者惊愕, 并且使得这个域 (更一般的情形是特征为 2 的域, 这是所有这样的域, 对于它 $1 + 1 = 0$ 成立, 其中符号 0 和 1 自然分别表示 K 的中性元和单位元, 参见 §30, 第 6 小节) 上的几何复杂化.

定理 1　*设 T 是域 K 上的一个向量空间, E 是相伴于 T 的一个仿射空间, 而 V 是 E 的一个非空子集. 则以下性质是等价的:*

a) V 是一个线性流形.

b) 对于每一个点 $P_0 \in V$, 所有点 $P \in V$ 对应的向量 $P - P_0$ 的集合是 T 的向量子空间.

c) 存在一个点 $P_0 \in V$, 使得所有点 $P \in V$ 对应的向量 $P - P_0$ 的集合是 T 的向量子空间.

如果这些等价的条件满足, 由向量 $P - P_0$ 组成的 T 的向量子空间不依赖 P_0 在 V 内的选取, 并且它是使得

$$s + P \in V \quad \text{对于所有 } P \in V$$

的 $s \in T$ 的集合.

我们证明 a) 蕴含 b). 为此只需证明对于任意 $Q, R \in V$ 和标量 $\lambda, \mu \in K$, 存在一个 $S \in V$, 使得

$$S - P_0 = \lambda(Q - P_0) + \mu(R - P_0),$$

而由这个关系定义的点 S 也满足

$$S - P_0 = \lambda(Q - P_0) + \mu(R - P_0) + (1 - \lambda - \mu)(P_0 - P_0),$$

因此有

$$S = \lambda Q + \mu R + (1 - \lambda - \mu)P_0.$$

如果 V 是线性流形, 那么这个 $S \in V$ 并且满足前面的关系.

显然 b) 蕴含 c), 我们证明 c) 蕴含 a). 考虑点

$$P = \sum_{i=1}^{n} \lambda_i P_i, \text{ 其中 } \sum_{i=1}^{n} \lambda_i = 1.$$

如果 V 含有 $P_1, \cdots, P_n$, 我们应当证明 $P \in V$. 而我们有

$$P - P_0 = \sum_{i=1}^{n} \lambda_i (P_i - P_0),$$

V 含有 P_i, 条件 c) 表明对于某个 $P' \in V$ 有 $P - P_0 = P' - P_0$, 而这显然表示 $P = P'$, 故得 $P \in V$.

我们现在证明, 如果 V 是一个非空线性流形, 那么由所有 $P - P_0 (P \in V)$ 组成的 T 的向量子空间不依赖 P_0 在 V 内的选取.

事实上设 H_0 是 T 的以这种方式对应于 P_0 的向量子空间, 而 H_1 是对应于另一个 $P_1 \in V$ 的子空间, 则对于所有 $P \in V$, 我们有

$$P - P_0 = (P - P_1) + (P_1 - P_0) = (P - P_1) - (P_0 - P_1).$$

由于根据定义, H_1 含有 $P - P_1$ 和 $P_0 - P_1$, 故 $P - P_0 \in H_1$. 这就表明 $H_0 \subset H_1$, 由对称性推理, $H_0 = H_1$.

我们发现由 $P - P_0 (P \in V)$ 组成的向量子空间不依赖 P_0 在 V 内的选取, 把它记作 H. 显然这也是形如 $P - Q$ 的向量的集合, 其中 P 和 Q 在 V 内变动. 设 $s \in H$ 和 $P \in V$, 选择 $P_0 = P$, 则存在 $Q \in V$, 使得 $s = Q - P$, 于是有 $Q = s + P$, 这就表明对于所有 $s \in H$ 和所有 $P \in V$ 有 $s + P \in V$.

反之, 考虑满足这个条件的一个 $s \in T$; 选择一个 $P \in V$, 并且令 $s + P = Q$, 则一方面 $Q \in V$, 另一方面 $s = P - Q$, 因此 $s \in H$, 这就完成了证明.

给定 E 内的一个非空线性流形 V, T 的满足 $P, Q \in V$ 的所有向量 $Q - P$ 组成的子空间 H 称为 V 的**主导子空间**.

我们知道 H 不足以确定 V, 不过如果两个非空线性流形 V' 和 V'' 有相同的主导子空间 H, 那么在 E 内存在一个平移把 V' 映射到 V''. 事实上, 选定 V' 的一个点 P' 和 V'' 的一个点 P'', 并考虑向量 $a = P'' - P'$. V 内的一个点 P 在 V' 内, 必须并且只需 $P - P' \in H$. 显然有

$$P - P' = (P + a) - P'',$$

由于 H 也是 V'' 的主导子空间, 我们发现 $P \in V'$ 和 $P + a \in V''$ 是等价的; 换句话说, 平移 $P \to P + a$ 映射 V' 到 V''.

例 12 设 T 是 K 上的向量空间, 像例 1 那样取集合 T 本身作为 E. 设 V 是一个线性流形, 而 H 是其主导子空间. 如果 $P_0 \in V$, 我们发现 V 是形式为所有 $P_0 + P$ 的向量 (不区分点和向量) 的集合, 其中 P 跑遍 H, 即 H 是 T 内包含 P_0 的陪集 (§7, 第 6 小节).

鉴于有些作者把我们所称的向量子空间称为线性流形, 在把向量空间看作仿射空间时, 把我们所称的线性流形称为**仿射线性流形**是合适的. 在一个向量空间 T 内, 一个仿射线性流形或者是空集, 或者是 T 的向量子空间的一个陪集.



注 2 设 V 是仿射空间 E 内的一个非空线性流形, 这里 E 相伴于一个向量空间 T, 设 $H \subset T$ 是 V 的主导子空间. 对于所有 $s \in H$ 和所有 $P \in V$ 有 $s + P \in V$, 于是得到从 $H \times V$ 到 V 内的一个映射 $(s, P) \to s + P$. 立刻发现这个映射满足仿射空间的公理 (EA1), (EA2) 和 (EA3). 因此, 可以把 V 考虑为相伴于 H 的一个仿射空间, 并且可以把仿射空间理论的所有定义应用到 V. 例如, 给定点 $P_i \in V$ 与其和为 1 的标量 $\lambda_i \in K$, 可以在仿射空间 V 内定义分配有质量 λ_i 的点 P_i 的质心, 这样获得的点显然跟在仿射空间 E 内所获得的是一样的.

注 3 假定 V 是由 E 中的点 $P_1, \cdots, P_n$ 生成的线性流形, 主导子空间 H 是形如 $P - M$ 的向量的集合, 其中 P 在 V 内变动, 而 M 在 V 内选定不再改变. V 的元素正是点

$$P = \sum_{0 \leqslant i \leqslant n} \xi_i P_i, \text{ 其中 } \sum_{0 \leqslant i \leqslant n} \xi_i = 1.$$

根据关系 (4), 对于这样的点有

$$P - M = \sum_{0 \leqslant i \leqslant n} \xi_i (P_i - M).$$

在特殊情形, 如果取 $M = P_0$, 我们发现 H 是向量

$$\xi_1(P_1 - P_0) + \cdots + \xi_n(P_n - P_0)$$

的集合, 其中标量 $\xi_1, \cdots, \xi_n$ 是任意的 (因为关系

$$\xi_0 + \xi_1 + \cdots + \xi_n = 1$$

显然允许从 $n + 1$ 个标量 $\xi_0, \cdots, \xi_n$ 中任意选取其中的 n 个). 即 H 是 T 的由向量

$$P_i - P_0 \quad (1 \leqslant i \leqslant n)$$

生成的子空间.

注 4 设 E 是一个仿射空间. 我们在第 2 小节曾经看到如果在 E 内选择一个 "原点" O, 那么可以把 E 看作一个向量空间, O 是其中的中性元, E 的两个元素和 $P + Q = R$ 的定义是

$$R - O = (P - O) + (Q - O).$$

用 E_O 表示这样得到的空间. 设 V 是 E 的非空子集, 并且取 V 的一个点作为 O, 那么 V 是 E 的一个线性流形, 必须并且只需 V 是 E_O 的一个向量子空间. 留给读者作为习题仔细证明.

5. 由直线生成线性流形

在初等几何里，用下列条件定义平面：如果它含有两个点，则它包含通过这两个点的直线. 由于我们采用了线性流形 (平面是其特殊情形) 的另一个定义，可以猜想，在这里采用的观点下，初等的"定义"要变成一个定理 (这就修正了这样一个事实，在古典几何里，我们的关于平面的定义事实上是一个定理). 只要 2 不是基础域 K 的特征 (即只要在 K 内 $1+1=0$ 不成立).

定理 2 设 T 是不以 2 为特征的域 K 上的一个向量空间，E 是相伴于 T 的一个仿射空间，而 V 是 E 的一个子集. V 是 E 的一个线性流形，必须并且只需对于不同的任意两个点 $P, Q \in V$，连接 P 和 Q 的直线包含于 V 内.

假定 V 是一个线性流形. 如果 V 含有不同的两个点 P 和 Q，则它也包含由 P 和 Q 生成的线性流形 (例 10)，即连接 P 和 Q 的直线.

反之，设这个条件满足. 由于空集是一个线性流形，我们可以假定 V 非空. 在选定一个点 $P_0 \in V$ 后，所有事情归结为 (定理 1) 指出对于 $Q \in V$，向量 $Q-P_0$ 的集合是 T 的一个向量子空间，即对于任意标量 $\lambda, \mu \in K$ 和点 $Q, R \in V$，存在 $S \in V$，使得

$$S - P_0 = \lambda(Q - P_0) + \mu(R - P_0).$$

但是由这个关系定义的 S 正是

$$S = \lambda Q + \mu R + (1-\lambda-\mu)P_0,$$

因此，变更记号，我们被引导至证明对于任意 $P, Q, R \in V$ 和标量 $\lambda, \mu, \nu \in K$，只要

$$\lambda + \mu + \nu = 1, \tag{11}$$

就有

$$\lambda Q + \mu R + \nu P \in V.$$

由于 K 不以 2 为特征，我们有 $2 \neq 0$，从而在 K 内 $3 \neq 1$，因此不可能有 $\lambda = \mu = \nu = 1$. 举例说可以假定 $\lambda \neq 1$. 那么 $1-\lambda$ 是可逆的，第 4 小节的公式 (10) 允许我们写出

$$\lambda Q + \mu R + \nu P = \lambda Q + (1-\lambda)M, \text{ 其中 } M = \frac{\mu}{1-\lambda}Q + \frac{\nu}{1-\lambda}R.$$

如果 $Q=R$，那么 $M=Q=R \in M$；如果 $Q \neq R$，点 M 位于连接 Q 和 R 的直线上，根据假设属于 V. 于是在所有情形都有 $M \in V$. 同样的推理表明 V 含有点 $\lambda Q + (1-\lambda)M$，这就完成了证明.

如果 K 是模 2 整数的域，定理的结论显然是错误的，因为在这个情形，E 的任意子集，只要它含有两个不同的点 P 和 Q，那么它包含连接 P 和 Q 的直线就没有任何特别之处，因为这条直线缩减为两个点本身 (例 11).

6. 有限维仿射空间, 仿射基

设 T 是域 K 上的一个向量空间, 而 E 是相伴于 T 的一个仿射空间. 我们说 E 是**有限维的**, 如果 T 是有限维的; 称数

$$\dim(\mathrm{E}) = \dim(\mathrm{T})$$

为 (基础域 K 上的) E 的**维数**.

假定 E 的维数是 n. 选择 T 的一个基 $(a_1, \cdots, a_n)$, E 的一个点 P_0, 并且令

$$\mathrm{P}_i = a_i + \mathrm{P}_0 \quad (1 \leqslant i \leqslant n).$$

对于所有 $\mathrm{P} \in \mathrm{E}$, 存在唯一的组标量 $\xi_1, \cdots, \xi_n \in \mathrm{K}$, 使得有

$$\mathrm{P} - \mathrm{P}_0 = \xi_1 a_1 + \cdots + \xi_n a_n.$$

令

$$\xi_0 = 1 - \xi_1 - \cdots - \xi_n,$$

那么显然上述关系可以写成

$$\mathrm{P} = \xi_0 \mathrm{P}_0 + \xi_1 \mathrm{P}_1 + \cdots + \xi_n \mathrm{P}_n, \tag{12}$$

我们立即发现对于所有 $\mathrm{P} \in \mathrm{E}$, *存在唯一的一组标量* $\xi_1, \cdots, \xi_n \in \mathrm{K}$ 满足 (12) 和

$$\xi_0 + \xi_1 + \cdots + \xi_n = 1. \tag{13}$$

反之, 为了使这个性质成立, 必须 $\mathrm{P}_i - \mathrm{P}_0 (1 \leqslant i \leqslant n)$ 组成 T 的一个基.

如果处于方才描述的情形, 则说族 $(\mathrm{P}_0, \cdots, \mathrm{P}_n)$ 是 E 的一个**仿射基**, 而满足 (12) 和 (13) 的标量 ξ_i 称为点 P 关于这个仿射基的**仿射坐标**.

例 13 取初等几何的通常空间作为 E, 一个仿射基是不在一个平面上的四个点 $(\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D})$ 的一个组, 即一个适当的四面体的四个顶点组成的一个组. 一个点 $\mathrm{P} \in \mathrm{E}$ 的仿射坐标是这样的数 x, y, z, t, 使得有

$$\mathrm{P} = x\mathrm{A} + y\mathrm{B} + z\mathrm{C} + t\mathrm{D}, \quad x + y + z + t = 1.$$

那么显然有

$$\overrightarrow{\mathrm{AP}} = y\overrightarrow{\mathrm{AB}} + z\overrightarrow{\mathrm{AC}} + t\overrightarrow{\mathrm{AD}},$$

于是 y, z, t 是向量 $\overrightarrow{\mathrm{AP}}$ 关于三个线性无关的向量 $\overrightarrow{\mathrm{AB}}, \overrightarrow{\mathrm{AC}}, \overrightarrow{\mathrm{AD}}$ 的分量, 这三个向量组成向量空间 $\mathrm{E_A}$ 的一个基.



注 5 设 V 是 E 内的一个非空线性流形, 我们已经看到 (注 2) 可以把 V 看作相伴于其主导子空间 H 的一个仿射空间. 称 H 的维数为**线性流形 V 的维数**. 如果 V 是由点 $P_0, \cdots, P_n$ 生成的, 注 3 表明 V 的维数等于向量组 $P_i - P_0 (1 \leqslant i \leqslant n)$ 的秩.

给定 E 内的两个线性流形 V 和 W, 如果 $V \subset W$, 则有 $\dim(V) \leqslant \dim(W)$, 仅当 $V = W$ 时才有 $\dim(V) = \dim(W)$. 留给读者作为习题仔细验证这个结果.

设 E 是相伴于有限维向量空间 T 的一个仿射空间. 上面我们已经看到, 为了点 $P_0, \cdots, P_n \in E$ 组成 T 的一个基, 只需 $P_i - P_0 (1 \leqslant i \leqslant n)$ 组成 T 的一个基. 由此推知 E 的所有仿射基含有同样数目的点, 即 $n+1$ 个, 这里 $n = \dim(E)$.

另外, 数目为 $n = \dim(E)$ 的向量 $P_i - P_0$ 组成 T 的一个基, 必须并且只需 (§19, 定理 10) 它们生成 T, 而这显然意味着所有 $P \in E$ 至少可以以一种方式写成

$$P = \sum_{0 \leqslant i \leqslant n} \xi_i P_i, \text{ 其中 } \sum_{0 \leqslant i \leqslant n} \xi_i = 1,$$

这还表明由 P_i 生成的线性流形是整个 E. 因为这个线性流形是包含 P_i 的最小的线性流形. 因此有下列结果:

定理 3 设 E 是域 K 上的维数为 n 的一个仿射空间. $n+1$ 个点组成 E 的一个仿射基, 必须并且只需它们不含于任何异于 E 本身的线性流形内.

例如, 为了四个点组成通常空间的一个仿射基, 它们不含于同一个平面是必要并且充分的.

注 6 设 $Q_0, \cdots, Q_m$ 是 E 的点. 说它们是线性无关的, 如果它们组成它们生成的线性流形 V 的一个基, 换句话说, 如果所有的点 $P \in E$ 至多有一种方式写成

$$P = \sum_{0 \leqslant i \leqslant m} \xi_i Q_i, \text{ 其中 } \sum_{0 \leqslant i \leqslant n} \xi_i = 1.$$

这显然也等于说向量 $Q_i - Q_0 (1 \leqslant i \leqslant m)$ 是线性无关的.

例如, 通常空间的三个点是线性无关的, 必须并且只需它们不位于同一条直线上, 即它们实际上生成一个平面.

7. 线性流形维数的计算

在实际中经常要计算由仿射空间的点的集合生成的线性流形的维数. 那么我们利用这里的结果:

定理 4 设 E 是域上的一个仿射空间, $P_0, \cdots, P_m$ 是 E 的一个仿射基, 并且

$$Q_j = \sum_{i=0}^{m} \alpha_{ij} P_i \quad (0 \leqslant j \leqslant n)$$

是 E 的 $n+1$ 个点. 设 r 是点 $Q_0, \cdots, Q_n$ 生成的线性流形的维数, 那么矩阵

$$A = (\alpha_{ij})_{0 \leqslant i \leqslant m, 0 \leqslant j \leqslant n}$$

的秩是 $r+1$.

根据注 5, 要找的秩等于向量组 $Q_j - Q_0 (1 \leqslant j \leqslant n)$ 的秩. 而我们有

$$\begin{aligned} Q_j - Q_0 &= (Q_j - P_0) - (Q_0 - P_0) = \sum_i \alpha_{ij}(P_i - P_0) - \sum_i \alpha_{i0}(P_i - P_0) \\ &= \sum_{i=1}^{m} (\alpha_{ij} - \alpha_{i0}) \cdot a_i, \end{aligned}$$

其中向量 $a_i = P_i - P_0 (1 \leqslant i \leqslant m)$ 组成 T 的一个基. 因此 (§19, 定理 5) 数 r 等于矩阵

$$A' = (\alpha_{ij} - \alpha_{i0})_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n}$$

的秩. 剩下的是利用点 Q_j 的仿射坐标之间的关系

$$\sum_{i=0}^{m} \alpha_{ij} = 1 \quad (0 \leqslant j \leqslant n) \tag{14}$$

证明矩阵 A 的秩比矩阵 A′ 的秩大 1.

为此, 一方面考虑齐次线性方程组

$$\sum_{j=0}^{n} \alpha_{ij}\xi_j = 0 \quad (0 \leqslant i \leqslant m), \tag{15}$$

另一方面考虑齐次线性方程组

$$\sum_{j=1}^{n} (\alpha_{ij} - \alpha_{i0})\eta_j = 0 \quad (0 \leqslant i \leqslant m). \tag{16}$$

第一个方程组的解组成 K^{m+1} 的一个子空间 M, 而第二个方程组的解组成 K^m 的一个子空间 N. 矩阵 A′ 的秩为 r, 根据 §19 定理 17, 我们有

$$\dim(N) = m - r.$$

如果暂时用 s 表示 A 的秩, 则有

$$\dim(M) = m + 1 - s.$$

为了证明 $s = r + 1$, 只需指出 M 和 N 有同样的维数, 即它们是同构的.

把 (15) 中的各个方程两端分别相加, 并且考虑到 (14), 显然得到

$$\xi_0 + \xi_1 + \cdots + \xi_n = 0,$$

因此 (15) 蕴含

$$\sum_{j=1}^{n} \alpha_{ij}\xi_j - \alpha_{i0}(\xi_1 + \cdots + \xi_n) = 0\,,$$

即

$$\sum_{j=1}^{n} (\alpha_{ij} - \alpha_{i0})\xi_j = 0.$$

于是令

$$u(\xi_0, \xi_1, \cdots, \xi_n) = (\xi_1, \cdots, \xi_n)$$

得到从 M 到 N 内的一个映射 u. 一个类似的推理表明令

$$v(\eta_1, \cdots, \eta_n) = (-\eta_1 - \cdots - \eta_n, \eta_1, \cdots, \eta_n)$$

可以构造从 N 到 M 内的映射 v. 显然 u 和 v 是从 M 到 N 内和从 N 到 M 内互逆的映射, 即从 M 到 N 上和从 N 到 M 上的同构, 而这就完成了定理的证明.



注 7 如果在 A 的第一列上加上其余各列的和, A 的秩不变, 这时第一列为 $(1, 1, \cdots, 1)$. 在这样得到的矩阵里, 从第一列后面各列减去第一列, 我们得到有同样秩的一个矩阵, 其第一行是 $(1, 0, \cdots, 0)$, 而第一行下方第一列右方的元素正是 A′ 的元素. 由此得到 $\mathrm{rg}(\mathrm{A}) = 1 + \mathrm{rg}(\mathrm{A}')$ ($\mathrm{rg}(\mathrm{A})$ 表示矩阵 A 的秩) 这个事实的另一个证明.

8. 仿射坐标下线性流形的方程

设 E 是交换域 K 上的 n 维仿射空间, 而 $(\mathrm{P}_0, \cdots, \mathrm{P}_n)$ 是 E 的一个仿射基. 考虑 E 内的 r 个无关的点

$$\mathrm{Q}_j = \sum_{i=0}^{n} \alpha_{ij}\mathrm{P}_i \quad (1 \leqslant j \leqslant r),$$

设 V 是它们生成的 $r-1$ 维线性流形. 设

$$\mathrm{P} = \sum_{i=0}^{n} \xi_i \mathrm{P}_i$$

是 E 的一个点. P 属于 V, 必须并且只需由 $(\mathrm{P}, \mathrm{Q}_1, \cdots, \mathrm{Q}_r)$ 生成的线性流形 W 仍然是 $r-1$ 维的. 事实上, 如果 $\mathrm{P} \in \mathrm{V}$, 显然有 $\mathrm{W} = \mathrm{V}$, 由此得到条件的必要性; 反之, 假定这个条件满足, 由于无论如何都有 $\mathrm{V} \subset \mathrm{W}$, 关系 $\dim(\mathrm{V}) = \dim(\mathrm{W})$ 表明 $\mathrm{V} = \mathrm{W}$, 故 P 属于 V.

根据定理 4, 为了 $\mathrm{P}\in\mathrm{V}$, 必须而且只需矩阵

$$\begin{pmatrix} \xi_0 & \alpha_{01} & \cdots & \alpha_{0r} \\ \xi_1 & \alpha_{11} & \cdots & \alpha_{1r} \\ \vdots & \vdots & & \vdots \\ \xi_n & \alpha_{n1} & \cdots & \alpha_{nr} \end{pmatrix}$$

的秩是 r. 由于 K 是交换的, 这也就是说从这个矩阵抽出的其阶 $s>r$ 的所有子方阵的行列式均为零 (§23, 第 8 小节, 注 5), 显然为此只需对于 $s=r+1$ 检查这个条件.

在检查 $s=r+1$ 所提及的条件时, 我们发现应当有

$$\begin{vmatrix} \xi_{i_0} & \alpha_{i_0 1} & \cdots & \alpha_{i_0 r} \\ \xi_{i_1} & \alpha_{i_1 1} & \cdots & \alpha_{i_1 r} \\ \vdots & \vdots & & \vdots \\ \xi_{i_r} & \alpha_{i_r 1} & \cdots & \alpha_{i_r r} \end{vmatrix} = 0 \quad 对于\ 0\leqslant i_0<i_1<\cdots<i_r\leqslant n; \tag{17}$$

刻画了 V 的点 P 的特征的这个关系称为**线性流形** V 关于所考虑的仿射基的方程.

最简单的情形是 $r=n$, 这时 V 的维数是 $n-1$ (这时称 V 是 E 内的一个**超平面**). 关系 (17) 缩减为仅有一个方程

$$\begin{vmatrix} \xi_0 & \alpha_{01} & \cdots & \alpha_{0n} \\ \xi_1 & \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \xi_n & \alpha_{n1} & \cdots & \alpha_{nn} \end{vmatrix} = 0.$$

由于在第一行加上其余各行的和行列式的值不变 (§24, 第 1 小节, 性质 e)), 并且一个点的仿射坐标之和总是 1, 我们发现可以用方程

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \xi_1 & \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \xi_n & \alpha_{n1} & \cdots & \alpha_{nn} \end{vmatrix} = 0 \tag{18}$$

代替前面的方程. 我们称 (18) 为 (无关的) 点 $\mathrm{Q}_1,\cdots,\mathrm{Q}_n$ 生成的**超平面的方程**.

第六章　多项式和代数方程

前面几章所发展的方法主要是用来研究线性方程. 但是众所周知, 同样重要并且复杂得多的是研究代数方程组, 即其左端是所考虑的未知元的单项式的线性组合. 这类方程组的一般研究是代数的基本对象之一, 并且引导到现在是数学中最活跃的分支之一的代数几何.

本章目的是引进十分初等并且无论如何在整个数学中都不可或缺的概念, 诸如代数关系、多项式、有理分式、代数方程等.

跟前面各节一样, 我们尽量在最大程度的一般性上引进这些概念. 举例来说, 我们定义系数在一个任意交换环 K 内的多变量多项式. 这里的一般化虽然不是漂亮的, 却是不加重负担的; 系数在一个环内而不是在一个域内的多项式概念是必需的, 如果我们期望能够在 n 个变量的多项式理论中对于 n 进行归纳推理 (n 个变量的多项式是系数在一个 $n-1$ 个变量的多项式环内的一个变量的多项式, 这个多项式环不是一个域). 自然还有更深刻的理由让我们采用如此高水平的一般化; 且不说这些, 即使假定基础环比如说就是实数域, 实际上多项式理论的叙述也不会有多少简化. (我们可以避开区分 “多项式” 和 “多项式函数”, 但是不得不证明 “恒等于零” 的多项式函数的系数全是零.)

注意 §26 至 §33 各节可以由仅学过 §0 至 §12 的读者阅读, §26 的 第 3 小节除外, 该小节的内容本来也不是针对初学者的.

§26 代数关系

1. 环的元素上的单项式和多项式

设 L 是一个交换环. 给定 L 的一个子环 K 和一个子集 B, 考虑同时包含 K 和 B 的 L 的子环. 这些子环的交集还是一个包含 K 和 B 的 L 的子环, 而且显然是具有这个性质的最小子环. 我们说这是**由 K 和 B 生成的 L 的子环**, 并且用记号

$$\mathrm{K}[\mathrm{B}]$$

表示它. 比如考虑这样一个情形, B 具有 n 个元素 $x_1, \cdots, x_n$, 这时改用记号

$$\mathrm{K}[x_1, \cdots, x_n]$$

表示 L 的由 K 和诸 x_i 生成的子环. 显然这个子环含有 $x_1, \cdots, x_n$ 的单项式, 即可以表示成

$$x_1^{r_1} \cdots x_n^{r_n}$$

的 L 的所有元素, 其中的指数是正整数或零. L 还含有单项式和 K 的一个元素的乘积, 从而含有所有有限个这种类型乘积的和, 即所有**系数在 K 内的 $x_1, \cdots, x_n$ 的多项式**. 系数在 K 内的 $x_1, \cdots, x_n$ 的多项式 $y \in \mathrm{L}$ 可以表示成形式

$$y = \sum_{r_1, \cdots, r_n \geqslant 0} a_{r_1 \cdots r_n} x_1^{r_1} \cdots x_n^{r_n}, \tag{1}$$

其中的系数 $a_{r_1 \cdots r_n} \in \mathrm{K}$ 几乎都是零(*). 如果愿意的话这样的多项式还可以写成形式

$$y = a + \sum_{1 \leqslant i \leqslant n} a_i x_i + \sum_{1 \leqslant i,j \leqslant n} a_{ij} x_i x_j + \sum_{1 \leqslant i,j,k \leqslant n} a_{ijk} x_i x_j x_k + \cdots \tag{1'}$$

其中的 “系数” $a, a_i, a_{ij}, a_{ijk}, \cdots$ 在 K 内并且几乎都为零. 为了以这种形式表示多项式, 只需在 (1) 先取满足 $r_1 + \cdots + r_n = 0$ 的项, 再取满足 $r_1 + \cdots + r_n = 1$ 的项, 再取满足 $r_1 + \cdots + r_n = 2$ 的项, 如此继续.

不仅环 $\mathrm{K}[x_1, \cdots, x_n]$ 含有所有的多项式 (1), 而且反过来, 这个环的所有元素都是这样的多项式, 即 $\mathrm{K}[x_1, \cdots, x_n]$ 刚好是可以写成形式 (1) 的所有 $y \in \mathrm{L}$ 的集合. 事实上, x_i 的两个单项式的乘积还是 x_i 的单项式, 由此直接发现多项式 (1) 的集合 L′ 是的 L 的含有 K 和 $x_i (1 \leqslant i \leqslant n)$ 的一个子环, 故包含 $\mathrm{K}[x_1, \cdots, x_n]$; 然而根据前面所说还有反向的包含关系, 故如所宣布的那样有 $\mathrm{L}' = \mathrm{K}[x_1, \cdots, x_n]$.

最简单的情形是 $n = 1$, 即 B 仅有一个元素 x. 系数在 K 内的 x 的多项式是这样的 $y \in \mathrm{L}$, 对于它存在一个整数 $p \geqslant 0$ 和 $a_0, \cdots, a_p \in \mathrm{K}$, 使得有

$$y = a_0 + a_1 x + \cdots + a_p x^p,$$

(*) 这表示 (§11, 第 5 小节) 使得 $a_{r_1 \cdots r_n} \neq 0$ 的族 $(r_1, \cdots, r_n)$ 只有有限个. 为了 (1) 的右端作为 L 的元素的和有意义, 这个条件是必要的.

它们组成由 K 和 x 生成的子环 K$[x]$.

注 1 不管对于多项式已有的成见, 初学者应当注意到字母 x 或 $x_1, \cdots, x_n$ 表示环 L 的固定元素, 而非 "变量".

例 1 我们有 $\mathbf{C} = \mathbf{R}[i]$, 因为所有复数写成形式 $a + bi$, 其中的 $a, b \in \mathbf{R}$.

例 2 取 $\mathrm{K} = \mathbf{Q}, \mathrm{L} = \mathbf{R}$, 并考虑 L 的子环 K$[x]$, 其中

$$x = \sqrt[3]{2}.$$

x 的逐次幂是

$$1, x, x^2, 2, 2x, 2x^2, 4, 4x, 4x^2, 8, \cdots,$$

因此 $\mathbf{R}$ 的子环 $\mathbf{Q}[x]$ 是可以表示成形式

$$a + bx + cx^2, \quad \text{其中 } a, b, c \in \mathbf{Q}$$

的实数的集合.

例 3 设 K 是一个交换环, 而 L 是从 K 到 K 内的映射的环 (§8, 例 3). 令 K 的每个元素对应一个由

$$f(t) = a \quad \text{对于所有 } t \in \mathrm{K}$$

定义的常值函数, 可以把 K 看作 L 的子环. 用 x 表示从 K 到 K 内的由

$$x(t) = t \quad \text{对于所有 } t \in \mathrm{K}$$

定义的恒等映射. L 的子环 K$[x]$ 的元素称为**环 K 上的多项式函数**. 这些显然是从 K 到 K 内的映射 f, 对于它们存在一个整数 $r \geqslant 0$ 和 K 的元素 $a_0, \cdots, a_r$, 使得有

$$f(t) = a_0 + a_1 t + \cdots + a_r t^r \quad \text{对于所有 } t \in \mathrm{K}.$$

这个记号在 §28 将被推广.

2. 代数关系

设 L 是一个交换环, K 是 L 的一个子环, 而 $x_1, \cdots, x_n$ 是 L 的有限个元素. 称 $x_1, \cdots, x_n$ 的单项式之间的系数在 K 内的所有线性关系是**系数在 K 内的 $x_1, \cdots, x_n$ 之间的代数关系**, 即代数关系是 K 的几乎所有元素为零的族 $(a_{r_1 \cdots r_n})_{r_1, \cdots, r_n \geqslant 0}$, 使得

$$\sum_{r_1, \cdots, r_n \geqslant 0} a_{r_1 \cdots r_n} x_1^{r_1} \cdots x_n^{r_n} = 0. \tag{2}$$

如果 (2) 仅当所有系数是零时成立, 则称 $x_1, \cdots, x_n$ 是**环 K 上代数无关的**; 在相反情形, 即至少存在一个非平凡的关系 (2), 则称 $x_1, \cdots, x_n$ 是**环 K 上代数相关的**.

取 $n=1$ 这个特殊情形, 即仅有一个 $x \in \mathrm{L}$. 如果 x 是 K 上代数相关的, 即对于至少一个整数 $p \geqslant 0$ 和不全为零的系数 $a_i \in \mathrm{K}$ 存在一个形式为

$$a_0 + a_1 x + \cdots + a_p x^p = 0 \tag{3}$$

的关系, 则称 x 在 K **上是代数的**. 在相反的情形, 则称 x 在 K **上是超越的**.

例 4 复数 i 在 $\mathbf{R}$ 上甚至在 $\mathbf{Q}$ 上是代数的, 因为它满足关系

$$i^2 + 1 = 0.$$

例 5 取 $\mathrm{K} = \mathbf{Q}$, $\mathrm{L} = \mathbf{C}$ 和

$$x = \sqrt[3]{2 - \sqrt{3}},$$

那么 x 在 $\mathbf{Q}$ 上是代数的. 事实上有 $x^3 = 2 - \sqrt{3}$, 因此

$$(x^3 - 2)^2 = 3,$$

此式还可以改写为

$$x^6 - 4x^3 + 1 = 0.$$

复数域 $\mathbf{C}$ 上的元素如果在有理数域 $\mathbf{Q}$ 上是代数的, 则称为**代数数** (§11, 例 11). 这是满足至少一个形式为

$$a_r z^r + \cdots + a_1 z + a_0 = 0$$

的关系的 $z \in \mathbf{C}$, 其中诸 a_i 是不全为零的有理数 (此外消去分母后可以假定它们是整数).

例 6 不是代数数的复数称为**超越数**. Liouville 在 1844 年给出这类数的第一批例子. 在 1882 年, Lindemann 证明数 $\pi = 3.14159\cdots$ 是超越数, 这个结果蕴含所谓的 "化圆为方" 问题的不可解性, 这否定了自古以来大多数数学家的猜想.

同样还证明了 (Hermite, 1873) 分析中的数 $e = 2.71828\cdots$ 是超越数.

例 7 设 L 是从 $\mathbf{R}$ 到 $\mathbf{R}$ 内的所有映射的环, 而 K 是多项式函数组成的 L 的子环 (例 3), 即由常值函数和由对于所有的 $t \in \mathbf{R}$ 令 $x(t) = t$ 得到的函数生成的 L 的子环. 我们称 $f \in \mathrm{L}$ 是**代数的**, 如果元素 x 和 f 在 L 的子环 $\mathbf{R}$ 上是代数相关的, 即存在几乎全部为零的常数 a_{pq}, 使得有

$$\sum_{p,q \geqslant 0} a_{pq} t^p f(t)^q = 0 \quad \text{对于所有 } t \in \mathbf{R}. \tag{4}$$

举例说由

$$f(t) = \sqrt[3]{t^2 - 1}$$

给定的函数显然是代数的. 一个函数如果不是代数的, 就称为**超越的**. 比如函数 $f(t) = \sin t$ 就是这种函数. 事实上, 假定它满足关系 (4), 令

$$f_p(t) = \sum_{q \geqslant 0} a_{pq} \cdot \sin^q t \quad (p \geqslant 0),$$

上面提到的关系就写成

$$\sum_p f_p(t) \cdot t^p = 0,$$

考虑到正弦函数的周期性对于所有整数 n 还有

$$\sum_p f_p(t) \cdot (t + 2n\pi)^p = 0.$$

对于给定的 $t \in \mathbf{R}$, 多项式 $\sum f_p(t) \cdot x^p$ 对于 x 的无穷多个值为零, 这就像后面将要指出的那样 (§32, 第 4 小节) 蕴含它的系数 $f_p(t)$ 全是零. 于是关系 (4) 蕴含对于任意 $p \geqslant 0$ 和 $t \in \mathbf{R}$ 有

$$\sum_{q \geqslant 0} a_{pq} \cdot \sin^q t = 0.$$

但是这时对于所有整数 $p \geqslant 0$, 多项式 $\sum a_{pq} x^q$ 对于 x 的无穷多个值 (可以令 $x = \sin t$, 即对于所有 $x \in [-1, 1]$) 变为零, 于是它的所有系数为零, 我们终于发现对于任意 p 和 q 系数 $a_{pq} = 0$, 这就证明了正如所宣布的那样 $\sin t$ 是超越函数.

3. 交换域的情形

首先证明下列结果:

定理 1 设 L 是一个域, K 是 L 的一个子域, 而 x 是 L 的一个元素. 则以下性质是等价的:

a) x 在 K 上是代数的;

b) K$[x]$ 作为 K 上的向量空间是有限维的;

c) K$[x]$ 是 L 的一个子域.

假定 x 在 K 上是代数的, 那么有一个关系

$$c_0 + c_1 x + \cdots + c_p x^p = 0,$$

其中的系数 $c_i \in$ K 不全为零. 可以假定 $c_p \neq 0$, 由于 K 是一个交换域, 由前面的关系得到一个形式为

$$x^p = a_0 + a_1 x + \cdots + a_{p-1} x^{p-1} \tag{5}$$

的关系, 其中的系数

$$a_i = -c_i/c_p$$

在 K 内. 我们要由此推出更一般的结果, 即对于每个 $n \geqslant 0$, 单项式 x^n 是 K$[x]$ 的元素 $1, x, \cdots, x^{p-1}$ 的系数在 K 内的一个线性组合. 对于 $n = 0$ 这时显然的, 这就允许关于 n 进行归纳推理. 假定形如

$$x^{n-1} = d_0 + d_1 x + \cdots + d_{p-1} x^{p-1}$$

的关系成立, 其中的 $d_i \in$ K, 考虑到关系 (5),

$$\begin{aligned} x^n = x \cdot x^{n-1} &= d_0 x + \cdots + d_{p-2} x^{p-1} + d_{p-1} x^p + d_{p-1}(a_0 + a_1 x + \cdots + a_{p-1} x^{p-1}), \\ &= d_0 x + \cdots + d_{p-2} x^{p-1} \end{aligned}$$

这显然就证明了我们的断言. 于是我们发现了, 如果 x 是代数的, 则存在一个整数 p, 使得 x 的每一个幂都是系数在 K 内的幂

$$1, x, \cdots, x^{p-1}$$

的一个线性组合. 因此显然这 p 个元素生成 K 上的向量空间 K$[x]$, 这就证明了定理中所陈述的 a) 蕴含 b).

现在证明 b) 蕴含 c). 令 K$[x]$ = F, 对于一个非零的 $a \in$ F, 考虑由

$$f(u) = au \quad 对于所有\ u \in \mathrm{F}$$

定义的映射

$$f : \mathrm{F} \to \mathrm{F}.$$

把 F 看作 K 上的一个向量空间, 显然 f 是 F 的一个自同态, f 的核由使得 $au = 0$ 的 $u \in$ F 组成. 由于 F (作为域的子环) 是一个整环, 并且因为 $a \neq 0$, 我们发现 Ker(f) 缩减为 $\{0\}$, 故 f 是单射的. 如果 F 是有限维的, 那么 (§19, 定理 13 的推论 1) f 是满射的, 作为特殊情形, 存在一个 $u \in$ F, 使得 $au = 1$. 这就证明了 F 的每个非零元素在 F 内是可逆的, 故 F 是 L 的一个子域.

剩下的是证明 c) 蕴含 a). 如果 K$[x]$ 是 L 的一个子域, 则 K$[x]$ 的每个非零元的逆元在 K$[x]$ 内, 特别是 K$[x]$ 含有 x 的逆元, 于是有形式为

$$x^{-1} = a_0 + a_1 x + \cdots + a_r x^r$$

的关系, 其中的系数 $a_i \in$ K, 由于这个关系还可以写成

$$a_r x^{r+1} + \cdots + a_0 x - 1 = 0,$$

其中的系数 $a_j \in$ K, 因此 x 是代数的, 至此证明完成.

例 8 取 $\mathrm{K}=\mathbf{Q}$ 和 $\mathrm{L}=\mathbf{C}$. 设 x 是一个代数数, 即 $\mathbf{C}$ 的一个元素, 它满足形式为

$$a_n x^n + \cdots + a_1 x + a_0 = 0$$

的关系, 其中的系数 $a_0, \cdots, a_n$ 不全是零 (例 5). 那么可以写成带有理系数 $a_0, \cdots, a_n$ 的形式

$$c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$$

的复数的集合是 $\mathbf{C}$ 的一个子域. 把这个结果同 §8 的例 2 做比较.

定理 2 设 L 是一个交换域, K 是 L 的一个子域, 而 $\overline{\mathrm{K}}$ 是 K 上的代数元素的集合, 则 $\overline{\mathrm{K}}$ 是 L 的一个子域.

设 x 和 y 在 K 上是代数的. 那么存在整数 p 和 q, 使得 x (对应的, y) 的每一个幂是系数在 K 内的

$$1, x, \cdots, x^{p-1} \text{ (对应的 } 1, y, \cdots, y^{q-1})$$

的线性组合. 通过乘法, 显然得到所有单项式 $x^r y^s$ 是 pq 个单项式

$$x^i y^j \quad (0 \leqslant i \leqslant p-1, 0 \leqslant j \leqslant q-1)$$

的系数在 K 内的线性组合. 故 L 的子环 $\mathrm{K}[x,y]$ 在 K 上是有限维的. 对于所有 $z \in \mathrm{K}[x,y]$, 子环 $\mathrm{K}[z]$ 包含于 $\mathrm{K}[x,y]$ 内, 更加在 K 上是有限维的, 因此 $\mathrm{K}[x,y]$ 的所有元素在 K 上是代数的. 特别说来, $x-y$ 和 xy 在 K 上是代数的, 这就已经证明了 $\overline{\mathrm{K}}$ 是 L 的一个子环.

为了完成证明, 还要指出如果 $x \neq 0$ 是一个代数数, 那么 x^{-1} 也是代数数. 而如果有一个系数在 K 内的关系

$$a_n x^n + \cdots + a_1 x + a_0 = 0,$$

显然 $x^{-1} = y$ 满足

$$a_n + \cdots + a_1 y^{n-1} + a_0 y^n = 0,$$

这就完成了证明.

例 9 代数数的集合是 $\mathbf{C}$ 的一个子域.

愿意了解在 "高等" 代数中扮演重要角色的这些问题的读者, 在本节和以下各节的习题中, 以及书末的参考文献中的若干著作中 (Van der Waerden, Samuel-Zariski, Lang) 找到许多补充结果.

§26 习题

1. 证明下列复数是代数数, 并且对于它们中每一个复数组成一个有理系数的代数方程:

$$\sqrt{2}+\sqrt{3};\quad \sqrt[3]{2}+\sqrt{3};\quad \sqrt[4]{2}+\sqrt[3]{3};\quad \sqrt{2}+\sqrt{3}+\sqrt{5}.$$

2. 设 L 是一个交换域, K 是 L 的一个子域, 而 x 是 L 的一个在 K 上超越的元素. 找出在元素

$$x^2+1 \quad 和 \quad x^3$$

之间存在的系数在 K 内的所有代数关系. 对于

$$x^3+x+1 \quad 和 \quad x^5$$

解答同样的问题.

¶3. 设 A 是一个交换整环, 而 K 是 A 的一个子环. 假定 A 最为作为 K 上的向量空间是有限维的. 证明 A 是一个域.

设 L 是一个交换域, K 是 L 的一个子域, 而 $x_1,\cdots,x_n$ 是 L 的在 K 上是代数的元素. 证明 L 的子环 $\mathrm{K}[x_1,\cdots,x_n]$ 是一个域.

¶¶4. 设 K 是一个交换域. 称所有允许 K 作为其子域的域 L 为 K 的**扩张** (例如 **C** 为 **R** 的一个扩张, 而 **R** 是 **Q** 的一个扩张). 那么就可以把 L 看作是 K 上的一个向量空间. 如果 L 在 K 上是有限维的, 即存在有限个元素 $a_1,\cdots,a_r\in \mathrm{L}$, 使得 L 的所有元素可以写成形式

$$x_1a_1+\cdots+x_ra_r,$$

其中的 $x_i\in \mathrm{K}(1\leqslant i\leqslant r)$, 则说 L 是 K 的一个**有限扩张**; 作为 K 上的向量空间的 L 的维数称为 L 在 K 上的**次数**, 记为

$$[\mathrm{L}:\mathrm{K}].$$

看作 K 上的向量空间 L 的所有的基称为 L 在 K 上的**基**. 当 $\mathrm{K}=\mathbf{Q}$, K 的有限扩张, 根据定义, 是**代数数域** [历史上, 曾要求代数数域是 **Q** 的包含于 **C** 内的有限扩张, 但是, 容易看出 **Q** 的所有有限扩张可以 "嵌入" 到 **C** 内, 以至于这个条件是多余的]. 下面用 K 表示一个交换域, 而用 L 表示 K 的一个 n 次扩张. 对于所有 $a\in \mathrm{L}$, 用 u_a 表示从 L 到 L 内的由

$$u_a(x)=ax \quad 对于所有\ x\in \mathrm{L}$$

给定的映射.

a) 证明 u_a 是看作 K 上的向量空间的 L 的自同态, 并且对于任意 $a,b\in \mathrm{L}$ 有关系

$$u_a+u_b=u_{a+b},\quad u_a\circ u_b=u_{a+b}.$$

和所有与 u_a 可交换的 (看作 K 上的向量空间的) L 的自同态是什么?

b) 对于所有 $a\in \mathrm{L}$ 有

$$\mathrm{Tr}_{\mathrm{L}/\mathrm{K}}(a)=\mathrm{Tr}(u_a),\quad \mathrm{N}_{\mathrm{L}/\mathrm{K}}(a)=\det(u_a)$$

(把 u_a 看作 K 上的有限维向量空间的自同态; u_a 的行列式在 §23 第 5 小节定义, 而迹在 §19 习题 22 定义). 称 $\mathrm{Tr}_{\mathrm{L/K}}(a)$ 和 $\mathrm{N}_{\mathrm{L/K}}(a)$ 分别是 a 的 (关于子域 K 的) **迹和模**, 它们当然是 K 的元素. 证明对于所有 $a, b \in \mathrm{L}$, 有

$$\mathrm{Tr}_{\mathrm{L/K}}(a+b) = \mathrm{Tr}_{\mathrm{L/K}}(a) + \mathrm{Tr}_{\mathrm{L/K}}(b), \quad \mathrm{N}_{\mathrm{L/K}}(ab) = \mathrm{N}_{\mathrm{L/K}}(a)\mathrm{N}_{\mathrm{L/K}}(b),$$

并且证明, 如果 L 在 K 上是 n 次的, 则对于所有的 $a \in \mathrm{K}$, 有

$$\mathrm{Tr}_{\mathrm{L/K}}(a) = na, \quad \mathrm{N}_{\mathrm{L/K}}(a) = a^n.$$

c) 设 $(a_i)_{1\leqslant i\leqslant n}$ 是看作 K 上的向量空间的 L 的一个基, 所有 $x \in \mathrm{L}$ 以唯一的方式写成形式

$$x = \xi_1 a_1 + \cdots + \xi_n a_n, \quad \text{其中 } \xi_1, \cdots, \xi_n \in \mathrm{K}.$$

令

$$x a_i = \sum_{1\leqslant j\leqslant n} \lambda_{ij} a_j,$$

其中 λ_{ij} 是 K 的元素. 用 λ_{ij} 计算 $\mathrm{Tr}_{\mathrm{L/K}}(x)$ 和 $\mathrm{N}_{\mathrm{L/K}}(x)$.

d) 设 K 的特征为 0 (即如果 $x \in \mathrm{K}$ 和 $\in \mathbf{Z}$ 满足 $rx = 0$, 则 $r = 0$ 或 $x = 0$; 参见 §30, 第 6 小节). 证明如果 $a \in \mathrm{L}$ 满足

$$\mathrm{Tr}_{\mathrm{L/K}}(ax) = 0 \quad \text{对于所有 } x \in \mathrm{L},$$

则有 $a = 0$. 由此推出, 如果 $(a_i)_{1\leqslant i\leqslant n}$ 是 L 在 K 上的一个基, 则有

$$\det(\mathrm{Tr}_{\mathrm{L/K}}(a_i a_j)_{1\leqslant i,j\leqslant n}) \neq 0.$$

e) 设 $(a_i)_{1\leqslant i\leqslant n}$ 是 L 在 K 上的一个基, 考虑 L 的 n 个元素

$$b_i = \sum_{1\leqslant j\leqslant n} \rho_{ij} a_j \quad (\rho_{ij} \in \mathrm{K}, 1 \leqslant i \leqslant n).$$

引进矩阵

$$\mathrm{A} = (\mathrm{Tr}_{\mathrm{L/K}}(a_i a_j))_{1\leqslant i,j\leqslant n},$$
$$\mathrm{B} = (\mathrm{Tr}_{\mathrm{L/K}}(b_i b_j))_{1\leqslant i,j\leqslant n},$$

证明

$$\det(\mathrm{B}) = \det(\mathrm{A})\det(\rho_{ij})^2.$$

由此推出 (如果 K 的特征为 0): L 的 n 个元素 $x_1, \cdots, x_n$ 组成在 K 上的 L 的一个基, 必须并且只需矩阵

$$(\mathrm{Tr}_{\mathrm{L/K}}(x_i x_j))_{1\leqslant i,j\leqslant n}$$

的行列式非零. 这个行列式称为 L 的元素 $x_1, \cdots, x_n$ 的**判别式**, 并且通常记作

$$\mathrm{D}_{\mathrm{L/K}}(x_1, \cdots, x_n).$$

f) 假定 K 的特征为 0, 设 $(u_i)_{1\leqslant i\leqslant n}$ 是 K 上的 L 的一个基. 证明存在 K 上的 L 的另一个基有 $(v_i)_{1\leqslant i\leqslant n}$, 使得

$$\mathrm{Tr}_{\mathrm{L/K}}(u_i v_j)=\begin{cases}1, & \text{如果 } i=j,\\ 0, & \text{如果 } i\neq j\end{cases}$$

(可以证明 v_i 关于基 (u_i) 的坐标由一个 Cramer 方程组给定). 称 (u_i) 和 (v_i) 是**互补的**.

g) 保留问题 f) 的假设和记号, 证明所有 $x\in\mathrm{L}$ 关于基 (v_i) 的坐标是 K 的元素

$$\mathrm{Tr}_{\mathrm{L/K}}(xu_i).$$

h) 不再假定 K 的特征是 0. 称 L 是 K 的一个**可分扩张**, 如果存在一个 $x\in\mathrm{L}$, 满足

$$\mathrm{Tr}_{\mathrm{L/K}}(x)\neq 0.$$

证明问题 d), e), f) 和 g) 的结果在这种情形仍然是有效的.

¶¶5. 设 L 是一个交换域, 而 K 是 L 的一个子域. 假定 L 是 K 的一个有限扩张 (习题 4), 并且用 E 表示包含 K 的 L 的一个子域.

a) 证明 L 是 E 的有限扩张, 并且 E 是 K 的有限扩张.

b) 设 $(a_i)_{1\leqslant i\leqslant r}$ 是 L 在 E 上的一个基, 而 $(b_j)_{1\leqslant j\leqslant s}$ 是 E 在 K 上的一个基. 证明 rs 个向量 a_ib_j 组成在 K 上 L 的一个基. 由此推出, 如果用 $[\mathrm{L}:\mathrm{K}]$ 表示 L 在 K 上的次数 (即作为 K 上的向量空间 L 的维数), 则有关系

$$[\mathrm{L}:\mathrm{K}]=[\mathrm{L}:\mathrm{E}][\mathrm{E}:\mathrm{K}].$$

c) 证明对于所有的 $x\in\mathrm{L}$ 有

$$\mathrm{Tr}_{\mathrm{L/K}}(x)=\mathrm{Tr}_{\mathrm{E/K}}(\mathrm{Tr}_{\mathrm{L/E}}(x)),$$
$$\mathrm{N}_{\mathrm{L/K}}(x)=\mathrm{N}_{\mathrm{E/K}}(\mathrm{N}_{\mathrm{L/E}}(x)).$$

(关于这里使用的记号参见习题 4.)

d) 设 x 是 L 的一个元素, 并且

$$x^s-a_{s-1}x^{s-1}+\cdots+(-1)^s a_0=0$$

是 x 所满足的系数在 K 内的次数最低的代数方程. 证明元素

$$1,x,\cdots,x^{s-1}$$

组成 K 上的域 $\mathrm{K}[x]$ 的一个基. 利用习题 4 的问题 c), 并且令 $\mathrm{K}[x]=\mathrm{E}$, 证明

$$\mathrm{Tr}_{\mathrm{E/K}}(x)=a_{s-1},\quad \mathrm{N}_{\mathrm{E/K}}(x)=a_0.$$

由此推出

$$\mathrm{Tr}_{\mathrm{L/K}}(x)=\frac{n}{s}a_{s-1},\quad \mathrm{N}_{\mathrm{E/K}}(x)=(a_0)^{n/s},$$

其中 $n=[\mathrm{L}:\mathrm{K}]$.

§27 多项式环

设 K 是一个交换环, 而 n 是一个正整数. 在本节我们打算构建一个交换环 L 和 n 个元素 $\mathrm{X}_1, \cdots, \mathrm{X}_n$, 使得下列条件满足:

(AP1) K 是 L 的一个子环;

(AP2) 元素 $\mathrm{X}_1, \cdots, \mathrm{X}_n$ 在 K 上是代数无关的;

(AP3) L 由 K 和 $\mathrm{X}_1, \cdots, \mathrm{X}_n$ 生成.

在研究一般情形之前, 我们在 $n = 1$ 这一特殊情形下解决这个问题, 然后构建一个包含 K 的环 L 和一个 K 上超越的 (§26, 第 2 小节) $\mathrm{X} \in \mathrm{L}$, 使得 $\mathrm{L} = \mathrm{K}[\mathrm{X}]$.

1. 一个未定元情形的预备知识

假定我们已经构建了 K 的一个母环 L 和一个 K 上超越的 $\mathrm{X} \in \mathrm{L}$, 使得 $\mathrm{L} = \mathrm{K}[\mathrm{X}]$, 那么每一个 $f \in \mathrm{L}$ 都可以用唯一的一种方式写成

$$f = a_0 + a_1\mathrm{X} + \cdots + a_n\mathrm{X}^n + \cdots, \tag{1}$$

其中 $a_n \in \mathrm{K}$ 几乎全部为零. 事实上, 由于 $\mathrm{L} = \mathrm{K}[\mathrm{X}]$, 幂

$$1, \mathrm{X}, \cdots, \mathrm{X}^n, \cdots$$

生成 K 上的模 L, 由于 X 在 K 上是超越的, 这些幂在 K 上是线性无关的 (即形成 §11 第 5 小节的意义下的 K 上的 L 的基). 反之, 显然 K 的每一个几乎全部元素为零的序列 $(a_n)_{n\geqslant 0}$ 由公式 (1) 定义 L 的一个元素.

设

$$f = a_0 + a_1\mathrm{X} + \cdots, \quad g = b_0 + b_1\mathrm{X} + \cdots$$

是 L 的两个元素, 令

$$f + g = c_0 + c_1\mathrm{X} + \cdots,$$
$$fg = d_0 + d_1\mathrm{X} + \cdots.$$

显然对于所有 $n \geqslant 0$ 有

$$c_n = a_n + b_n. \tag{2}$$

为了计算 fg, f 的每一项 $a_p\mathrm{X}^p$ 乘以 g 的每一项 $b_q\mathrm{X}^q$, 并且把所得的结果相加. 为了得到一项 X^n 必须取满足关系 $p + q = n$ 的 p, q, 并且这样的 p, q 对于计算 d_n 的贡献是 a_pb_q. 如此看来

$$d_n = \sum_{p+q=n} a_pb_q = a_nb_0 + a_{n-1}b_1 + \cdots + a_1b_{n-1} + a_0b_n. \tag{3}$$

公式 (2) 和 (3) 的益处在于用它们计算 $f + g$ 和 fg, 不涉及 X. 它们将是我们定义环 K[X] 的出发点.

2. 一个未定元的多项式

从 K 出发, 我们现在要构建满足所提条件的环 L (在本小节我们只是简单地定义 L 的元素和对于这些元素的代数运算, 待到下一小节再验证满足所提条件).

作为定义, L 将是 K 的几乎全为零的元素序列

$$f = (a_n)_{n \geqslant 0} = (a_0, a_1, \cdots)$$

的集合, 一个这样的序列称为**系数在 K 内的一个未定元的多项式**, 而元素 a_n 称为多项式 f 的**系数**.

给定一个系数在 K 内的多项式 $f = (a_n)_{n \geqslant 0}$, 称使得 $a_d \neq 0$ 的最大整数 d 为 f 的**次数**. 在对于所有 $n \geqslant 0$ 都有 $a_n = 0$ 这种情形, 这个定义失去意义. 在这种情形, 我们称符号 $-\infty$ 为 f 的次数 [这个符号自然不是一个普通的整数, 这是我们将使用的一个新的数学对象, 注意以下两个约定: 作为定义, 一方面令

$$-\infty + n = -\infty, \quad 如果\ n \in \mathbf{Z}\ 或\ n = -\infty,$$

而另一方面约定

$$-\infty \leqslant n, \quad 如果\ n \in \mathbf{Z}\ 或\ n = -\infty;$$


所有其他涉及符号 $-\infty$ 的运算, 比如

$$-\infty - (-\infty)$$

之类都被认为是没有意义的].

多项式 f 的次数用记号表示为

$$d^\circ(f).$$

在系数在 K 内的一个未定元的多项式的集合内, 将如下定义**加法**和**乘法**: 给定两个多项式

$$f = (a_n)_{n \geqslant 0}, \quad g = (b_n)_{n \geqslant 0},$$

多项式

$$f + g = (c_n)_{n \geqslant 0}, \quad fg = (d_n)_{n \geqslant 0}$$

由前一个小节的关系 (2) 和 (3) 给定:

$$c_n = a_n + b_n, \tag{2'}$$

$$d_n = a_n b_0 + a_{n-1} b_1 + \cdots + a_1 b_{n-1} + a_0 b_n = \sum_{r+s=n} a_r b_s. \tag{3'}$$

为了验证这个定义的合理性, 应当指出序列 (c_n) 和 (d_n) 还是多项式, 即当整数 n 充分大时有 $c_n = d_n = 0$. 设 p 和 q 分别是 f 和 g 的次数, 如果 $n > \max(p, q)$, 则有 $a_n = b_n = 0$, 故 $c_n = 0$, 这就证明了 $f + g$ 是多项式, 并且

$$d^\circ(f + g) \leqslant \max(d^\circ(f), d^\circ(g)). \tag{4}$$

再设 $n > p + q$, 那么在表达式 (3′) 的一般项 $a_r b_s$ 里, 或者 $r > p$ (因此 $a_r = 0$), 或者 $s > q$ (因此 $b_s = 0$), 故 (3′) 的这样的项全为零. 这个推理不仅表明 fg 也是多项式, 而且表明

$$d^\circ(fg) \leqslant d^\circ(f) + d^\circ(g). \tag{5}$$

我们现在应当验证配备了刚定义的运算的 L 是一个交换环. 为此要实施的计算原则上是平凡的, 难易程度跟 §9 第 3 小节的相当, 让读者细心验证大部分公理, 我们来验证乘法是结合的.

为此设

$$f = (a_n)_{n \geqslant 0}, \quad g = (b_n)_{n \geqslant 0}, \quad h = (c_n)_{n \geqslant 0}$$

是三个系数在 K 内的多项式. 令

$$fg = (u_n)_{n \geqslant 0}, \quad gh = (v_n)_{n \geqslant 0}.$$

$(fg)h$ 指标为 n 的系数等于

$$u_n c_0 + \cdots + u_0 c_n, \tag{6}$$

而 $f(gh)$ 指标为 n 的系数等于

$$a_n v_0 + \cdots + a_0 v_n, \tag{7}$$

于是一切归结为证明表达式 (6) 和 (7) 对于任意 n 是相等的. 而 u_p 是满足 $i + j = p$ 的乘积 $a_i b_j$ 的和, 因此 (6) 是以下表达式的和

$$(a_i b_j) c_k, \quad \text{其中 } (i + j) + k = n.$$

另外, v_q 是满足 $j + k = q$ 的乘积 $b_j c_k$ 的和, 故 (7) 是以下表达式的和

$$a_i (b_j c_k), \quad \text{其中 } i + (j + k) = n;$$

由这些显然得到 (6) 和 (7) 是相等的.

容易看出 L 的元素 0 和 1 是

$$0 = (0, 0, 0, \cdots),$$
$$1 = (1, 0, 0, \cdots).$$

3. 多项式记号

我们指出可以把 K 等同到 L 的一个子环. 为此, 对于每一个 $a \in \mathrm{K}$ 令

$$j(a) = (a, 0, 0, \cdots)$$

定义一个映射

$$j : \mathrm{K} \to \mathrm{L}.$$

显然 j 是单射的, 借助公式 (2) 和 (3) 可以简单地验证

$$j(a+b) = j(a) + j(b), \quad j(ab) = j(a)j(b),$$
$$j(0) = 0, \quad j(1) = 1.$$

因此 j 是从 K 到 L 的一个子环上的同构, 也就是说, 为了对于 K 的元素实施代数运算, 就用它们在 j 下的像代替它们, 并且对于这样得到的元素实施所提及的运算. 于是把 K 的一个元素 a 等同于 L 的对应元素就是顺理成章的事情了. 今后我们就把这种等同写成

$$a = (a, 0, 0, \cdots) \quad 对于每个\ a \in \mathrm{K}. \tag{8}$$

如果 K 的元素 $a \neq 0$, 那么这样得到的 L 的元素是系数在 K 内的 0 次多项式 (L 的对应 K 的元素 0 的元素的次数是 $-\infty$, 而非 0). L 的这些元素经常称为**常元**.

注 1 我们刚定义的 “常元” 概念是相对于基础环 K 而言: 这是系数在 K 内的零多项式或 0 次多项式, 如果愿意, 也就是 K 的一个元素 (看作系数在 K 内的多项式). 所用的术语下一节就会得到解释.

我们注意到, 由于 K 是 L 的一个子环, 可以把 L 看作 K 上的一个模, $a \in \mathrm{K}$ 和 $f \in \mathrm{L}$ 的乘积就是 f 和 L 的元素 (8) 的乘积. 利用 (3) 容易发现

$$a \cdot (b_0, b_1, \cdots) = (ab_0, ab_1, \cdots). \tag{9}$$

现在构造 L 的在 K 上超越的一个元素. 我们取

$$\mathrm{X} = (0, 1, 0, 0, \cdots). \tag{10}$$

利用 (3) 容易发现

$$\mathrm{X}^2 = (0, 0, 1, 0, 0, \cdots),$$
$$\mathrm{X}^3 = (0, 0, 0, 1, 0, \cdots),$$

等等, 由此利用 (9), 如果 $a_0, a_1, \cdots$ 是 K 的元素, 即得公式

$$\begin{aligned} a_0 &= (a_0, 0, 0, 0, \cdots), \\ a_1\mathrm{X} &= (0, a_1, 0, 0, \cdots), \\ a_2\mathrm{X}^2 &= (0, 0, a_2, 0, 0, \cdots), \\ a_3\mathrm{X}^3 &= (0, 0, 0, a_3, 0, \cdots), \end{aligned}$$

等等, 如果 a_i 几乎全部是零, 则有

$$a_0 + a_1\mathrm{X} + a_2\mathrm{X}^2 + \cdots = (a_0, a_1, a_2, \cdots). \tag{11}$$

考虑到

$$0 = (0, 0, 0, \cdots),$$

这个结果表明 (11) 的左端仅当 $a_0 = a_1 = a_2 = \cdots = 0$ 时为零, 因此, X 是 K 上的超越元.

另外, 关系 (11) 表明 L 的每一个元素都是系数在 K 内的一个多项式 (在 §26 第 1 小节的意义下), 换句话说

$$\mathrm{L} = \mathrm{K}[\mathrm{X}].$$

这就表明环 L 满足本节开头所说的条件.

在实际中, L 称为**系数在 K 内的一个未定元的多项式环**. 为了表示 L 的一个元素

$$f = (a_0, a_1, a_2, \cdots),$$

使用由关系 (11) 验证其合理性的记号

$$f = a_0 + a_1\mathrm{X} + a_2\mathrm{X}^2 + \cdots. \tag{12}$$

换句话说, 从现在开始, 读者可以忽略前面的小节, 它仅用来定义多项式. 对于后续的内容 (乃至数学中所有用到多项式的地方), 除了前面宣布的条件 (AP1), (AP2) 和 (AP3), 再不用记住任何东西了 —— 即, *系数在 K 内的一个多项式是这样一个数学对象, 可以用唯一的一种方式写成形式 (12), 其系数 $a_n \in \mathrm{K}$ 几乎全部为零, 并且把多项式看作一个交换环的元素对其进行计算, 在计算时可以利用 "显然的" 计算法则*.

注 2 读者或许会问为什么一开始不定义多项式为系数 (12) 的一个表达式, 对于它实施 "显然的" 加法和乘法法则的计算. 理由是: 在前面不能够给出现在 (12) 中的字母 X 以精确的数学涵义, 以致得到的多项式 "定义" 其实不是什么定义.





注 3　与有时还在流行的传统相反, 小心别把字母 X 当作代表 K 的 "变动元素"; 字母 X 表示特殊的多项式 (10), 它的定义没有什么任意性和不确定性.

字母 X 代表 K 的变动元素的想法来源于经常会犯的系数在 K 内的多项式和环 K 的多项式函数 (§26, 例 3) 之间的混淆. 这两个概念之间的关系将在下一节讨论.

不言而喻, 代替用 X 表示多项式 (10), 可以用任何其他的字母表示它 (在实际中经常使用字母 Y, Z, T 等), 只要所选择的字母没有为其他目的使用过.

4. 多个未定元的多项式

我们要用关于整数 n 的归纳法构造本节开头所提出的问题的解.

用 K′ 表示这样一个交换环, 它包含 K 作为子环, 并且由 K 和在 K 上代数无关的 $n-1$ 个元素 $X_1, \cdots, X_{n-1}$ 生成. 用 L 表示系数在 K′ 内的带一个未定元的多项式环, 并且用 X_n 表示所提到的未定元, 故有

$$L = K'[X_n] = K[X_1, \cdots, X_{n-1}][X_n].$$

我们要指出 L 符合条件 (AP1), (AP2) 和 (AP3).

由于 K 是 K′ 一个子环, 而后者本身是 L 的子环, (AP1) 的满足是平凡的.

设 L′ 是由 K 和 $X_1, \cdots, X_n$ 生成的 L 的子环, 由于它含有 K 和 $X_1, \cdots, X_{n-1}$, 因此它包含 K′. 由于它含有 K′ 和 X_n, 它包含 L, 从而等于 L. 因此 L 由 K 和 $X_1, \cdots, X_n$ 生成, 这就是条件 (AP3).

最后考虑关系

$$\sum a_{r_1 \cdots r_n} X_1^{r_1} \cdots X_n^{r_n} = 0,$$

其中 K 的元素 $a_{r_1 \cdots r_n}$ 几乎全是零. 引进 K′ 的元素

$$f_s = \sum_{r_1 \cdots r_{n-1}} a_{r_1 \cdots r_{n-1} s} X_1^{r_1} \cdots X_{n-1}^{r_{n-1}}.$$

所考虑的关系写成

$$\sum_s f_s X_n^s = 0,$$

由于 X_n 是 K′ 上超越的, 因此得到对于所有 $s \geqslant 0$ 有 $f_s = 0$. 而由于 $X_1, \cdots, X_{n-1}$ 在 K 上是代数无关的, 这就蕴含对于任意 $r_1, \cdots, r_{n-1}$ 和 s, 有

$$a_{r_1 \cdots r_{n-1} s} = 0,$$

这就验证了 (AP2).

我们刚定义的环

$$L = K[X_1, \cdots, X_{n-1}][X_n] = K[X_1, \cdots, X_n]$$

称为**系数在 K 内的 n 个未定元的多项式环**, 而它的元素称为**系数在 K 内的 n 个未定元的多项式**. 一个这样的元素按唯一的一种方式写成形式

$$f=\sum a_{r_1\cdots r_n}\mathrm{X}_1^{r_1}\cdots\mathrm{X}_n^{r_n}, \tag{13}$$

它带几乎全为零的**系数** $a_{r_1\cdots r_n}\in\mathrm{K}$. 在这些多项式上以 "显然的" 方式实施运算, 即把它们看作一个交换环的元素, 事实上它们本来就是这样.

举例说, 如果 $n=2$, 在这种情形经常以 X 和 Y 记未定元 X_1 和 X_2, 系数在 K 内的两个未定元的多项式是一个形如

$$f=a_{rs}\mathrm{X}^r\mathrm{Y}^s$$

的表达式, 带几乎全为零的系数 $a_{rs}\in\mathrm{K}$. 求和取遍所有非负整数偶 (r,s). 把 $r+s$ 具有一个给定值的项聚集在一起, 一个这样的多项式还可以写成形式

$$\begin{aligned}f=a_{00}&+(a_{10}\mathrm{X}+a_{01}\mathrm{Y})+(a_{20}\mathrm{X}^2+a_{11}\mathrm{XY}+a_{02}\mathrm{Y}^2)\\&+(a_{30}\mathrm{X}^3+a_{21}\mathrm{X}^2\mathrm{Y}+a_{12}\mathrm{XY}^2+a_{03}\mathrm{Y}^3)+\cdots,\end{aligned}$$

其中必然仅有有限个非零项.

同样, 如果 $n=3$, 用 X, Y, Z 表示未定元, 系数在 K 内三个未定元的多项式是一个带几乎全为零的系数 $a_{ijk}\in\mathrm{K}$ 的表达式

$$f=\sum_{i,j,k\geqslant 0}a_{ijk}\mathrm{X}^i\mathrm{Y}^j\mathrm{Z}^k,$$

一个这样的多项式还可以写成

$$\begin{aligned}f=a_{000}&+(a_{100}\mathrm{X}+a_{010}\mathrm{Y}+a_{001}\mathrm{Z})\\&+(a_{200}\mathrm{X}^2+a_{020}\mathrm{Y}^2+a_{002}\mathrm{Z}^2+a_{011}\mathrm{YZ}+a_{101}\mathrm{ZX}+a_{110}\mathrm{XY})+\cdots.\end{aligned}$$

不言而喻, 一个多项式的系数不必强求采用上面的记号: 在数学里, 每个人都可以自由选择他的记号 (只要它们是协调的), 比如, 代替用多重指标表示系数, 我们可以自由地使用不同的字母, 比如可以把两个未定元的多项式写成形式

$$f=a+b\mathrm{X}+c\mathrm{Y}+d\mathrm{X}^2+e\mathrm{XY}+f\mathrm{Y}^2+g\mathrm{X}^3+\cdots.$$

这种方法的不便之处在于拉丁字母仅提供有限种可能性. 另外, 希望读者立刻发现我们刚刚展示的记号中的不便之处.

5. 偏次数和总次数

如果在表达式 (13) 里把 X_i 的指数 r_i 取给定值的所有的项合并在一起, 就会发现可以把 f 看作一个未定元 X_i 其系数在环

$$\mathrm{K}_i=\mathrm{K}[\mathrm{X}_1,\cdots,\mathrm{X}_{i-1},\mathrm{X}_{i+1},\cdots,\mathrm{X}_n]$$

内的多项式. 考虑 f 作为多项式环 $\mathrm{K}_i[\mathrm{X}_i]$ 的元素, 可以定义 f 的次数, 称为 f **关于 X_i 的次数**. 可以定义:

$$f = \sum_{n \geqslant 0} u_n \mathrm{X}_i^n,$$

其中的 u_n 是系数在 K 内的 X_i 以外的未定元的多项式. 这样表示 f 之后, f 关于 X_i 的次数就是使得 $u_n \neq 0$ 的最大整数 n, 或是 $-\infty$, 如果 $f = 0$.

称存在满足条件

$$a_{r_1, \cdots, r_n} \neq 0, \quad r_1 + \cdots + r_n = d$$

的整数 $r_1, \cdots, r_n$ 的最大的整数 d 为 f 的**总次数**, 或简单地称为**次数**. 还可以如下定义总次数. 说一个系数在 K 内的多项式是 d **次齐次的**, 如果它实际含有的 (即系数非零的) 单项式的总次数都是 d. 在 (13) 中, 合并 $r_1 + \cdots + r_n$ 有给定值的项, 就会发现系数在 K 内的所有多项式都可以用唯一的一种方式写成形式

$$f = f_0 + f_1 + \cdots + f_r + \cdots,$$

其中的 f_r 对于所有 $r \geqslant 0$ 都是 r 次齐次的, 并且对于几乎全部 r 都是零. 这样表示之后, f 的总次数就是使得 $f_d \neq 0$ 的最大整数 d (总是约定多项式 0 的总次数为 $-\infty$).

f 的总次数跟仅一个未定元的多项式一样用记号 $d^\circ(f)$ 表示. 有不等式

$$d^\circ(f+g) \leqslant \max(d^\circ(f), d^\circ(g)),$$
$$d^\circ(fg) \leqslant d^\circ(f) + d^\circ(g).$$

事实上, 设 f 和 g 的次数分别是 p 和 q, 那么有

$$f = f_0 + \cdots + f_p, \quad g = g_0 + \cdots + g_q$$

(按照一般的方式, u_r 表示一个多项式的总次数为 r 的单项式的和). 当逐项相加时, 我们发现 $f+g$ 涉及不到次数超过 p, q 中较大者的单项式, 由此得到第一个不等式. 另外, fg 是乘积 $f_i g_j$ 的和; 但是显然 $f_i g_j$ 是总次数 $i+j$ 齐次的; 因此涉及乘积 fg 的单项式的总次数最多是 $p+q$, 这就证明了第二个不等式.

6. 系数在一个整环内的多项式

我们要证明

定理 1 设 K 是一个交换整环, 对于所有整数 n, 环 $\mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$ 是整环. 并且对于任意 $f, g \in \mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$ 有

$$d^\circ(fg) = d^\circ(f) + d^\circ(g).$$

为了证明 $K[X_1, \cdots, X_n]$ 是整环, 考虑关系

$$K[X_1, \cdots, X_n] = K[X_1, \cdots, X_{n-1}][X_n].$$

这使我们可以关于 n 进行归纳推理(*), 以便把问题归结为 $n = 1$ 的情形.

设

$$f = a_0 + \cdots + a_p X^p, \quad a_p \neq 0,$$
$$g = b_0 + \cdots + b_q X^q, \quad b_q \neq 0$$

是系数在 K 内的一个未定元的非零多项式. 显然 fg 仅含有一个次数为 $p+q$ 的项, 即

$$a_p b_q X^{p+q}.$$

由于 K 是整环, 我们有 $a_p b_q \neq 0$, 因此 $fg \neq 0$. 这就证明了定理的第一个断言.

为了证明定理的第二部分 (在多个未定元的情形) 写出

$$f = f_0 + \cdots + f_p, \quad f_p \neq 0,$$
$$g = g_0 + \cdots + g_q, \quad g_q \neq 0;$$

显然 fg 的总次数为 $p+q$ 的齐次部分是 $f_p g_q$. 而由于我们已经知道 $K[X_1, \cdots, X_n]$ 是一个整环, 故 $f_p g_q \neq 0$, 故 fg 的总次数是 $p+q$, 这就完成了证明.

注 4 认真说起来, 我们仅在 f 和 g 非零的情形下证明了关系



$$d^\circ(fg) = d^\circ(f) + d^\circ(g).$$

如果 $f = 0$, 这个关系写成 $-\infty = -\infty + d^\circ(g)$, 并且从第 2 小节中对于符号 $-\infty$ 所约定的计算法则推出.

注 5 显然如果 K 不是整环, $K[X_1, \cdots, X_n]$ 也不可能是整环. 此外, 在这种情形, 关系



$$d^\circ(fg) = d^\circ(f) + d^\circ(g)$$

可能也是错误的: 比如在 K 内选择两个非零元素 a 和 b, 使得 $ab = 0$, 取

$$f = aX, \quad g = bX,$$

则有

$$d^\circ(fg) = -\infty, \quad d^\circ(f) + d^\circ(g) = 2.$$

(本节习题见 §28 后.)

(*) 在多个未定元的多项式理论中经常这样处理 —— 但仅当允许系数在一个任意环内的多项式时这样做才是可能的, 因为即使 K 是一个域, 环 $K[X_1, \cdots, X_{n-1}]$ 也不是一个域.

§28 多项式函数

1. 多项式的值

设 K 是一个交换环, 而

$$f = \sum a_{r_1 \cdots r_n} \mathrm{X}_1^{r_1} \cdots \mathrm{X}_n^{r_n} \tag{1}$$

是系数在 K 内的 n 个未定元的一个多项式. 给定 K 的一个交换扩环 L, 用 L 的元素 $u_1, \cdots, u_n$ 代替表达式 (1) 中的字母 $\mathrm{X}_1, \cdots, \mathrm{X}_n$ 所得到的 L 的元素

$$\sum a_{r_1 \cdots r_n} u_1^{r_1} \cdots u_n^{r_n} \tag{2}$$

称为 f 在一个点 $u = (u_1, \cdots, u_n) \in \mathrm{L}^n$ 的**值**. f 在 u 的值可以记作下列记号中的一个:

$$f(u), \quad f(u_1, \cdots, u_n).$$

如果

$$f(u) = 0,$$

则称 u 是 f 的一个**零点**, 如果 $n = 1$, 还称 u 是 f 的一个**根**.

给定一个交换环 L 和 L 的一个子环 K, 所有的映射 $\varphi : \mathrm{L}^n \to \mathrm{L}$ 称为 L^n 上的**系数在 K 内的多项式函数**, 如果存在一个系数在 K 内的 n 个未定元的多项式 f, 使得

$$\varphi(u) = f(u) \quad 对于所有\ u \in \mathrm{L}^n.$$

更一般的, 称一个映射 $\varphi : \mathrm{L}^n \to \mathrm{L}^r$ 是系数在 K 内的**多项式**映射, 如果有 (§2, 第 9 小节)

$$\varphi = (\varphi_1, \cdots, \varphi_r),$$

其中的 φ_j 是 L^n 上的系数在 K 内的多项式函数.

例 1 如果取 K = L 和 $n = 1$, 我们显然重新得到 §26 例 3 的函数, 即从 K 到 K 内的函数, 其形式是

$$a_0 + a_1 t + \cdots + a_n t^n,$$

其中的 $a_0, \cdots, a_n$ 是 K 的 "固定" 元素, 而 t 表示 K 的 "变动" 元素.

例 2 K 和 n 是任意的, 系数在 K 内的一个多项式环取作 L, 即

$$\mathrm{L} = \mathrm{K}[\mathrm{Y}_1, \cdots, \mathrm{Y}_p].$$

对于所有多项式

$$f \in \mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$$

和任意

$$u_1, \cdots, u_n \in \mathrm{K}[\mathrm{Y}_1, \cdots, \mathrm{Y}_p]$$

可以定义 $f(u_1, \cdots, u_n)$; 基于显然的理由, 我们称这是**多项式 $u_1, \cdots, u_n$ 代换 f 的 $\mathrm{X}_1, \cdots, \mathrm{X}_n$ 得到的 $\mathrm{Y}_1, \cdots, \mathrm{Y}_p$ 的多项式**.

如果在特殊情形取 $\mathrm{L} = \mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$ 和

$$u_1 = \mathrm{X}_1, \cdots, u_n = \mathrm{X}_n,$$

这样得到的多项式 $f(u_1, \cdots, u_n)$ 显然就是 f 自己, 用记号

$$f = f(\mathrm{X}_1, \cdots, \mathrm{X}_n)$$

表示这个结果. 在实际中, 经常用记号 $f(\mathrm{X}_1, \cdots, \mathrm{X}_n)$ 代替 f. 这个记号的好处是使得在 f 中的未定元的记号一目了然, 但读者不应该忘掉字母 X_i 不表示 K 变动元素. 参见 §27 的注 3.

2. 多项式函数的和与乘积

设 K 是一个交换环, L 是包含 K 的一个交换环, 而 n 是至少等于 1 的一个整数. 给定 L 的元素 $u_1, \cdots, u_n$, 考虑由

$$v(f) = f(u_1, \cdots, u_n) \quad \text{对于所有多项式 } f \in \mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n],$$

给定的映射

$$v : \mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n] \to \mathrm{L}.$$

显然有

$$v(f) = f, \quad \text{如果 } f \in \mathrm{K} \text{ (即 } f \text{ 是一个 “常量”)}, \tag{3}$$

$$v(f) = u_i, \quad \text{如果 } f = \mathrm{X}_i. \tag{4}$$

另外, v 是环的同态. 即因为已经在 (3) 中指出 $v(1) = 1$, 对于任意 f 和 g 有关系

$$v(f + g) = v(f) + v(g), \tag{5}$$

$$v(fg) = v(f)v(g). \tag{6}$$

事实上, 令

$$f = \sum a_{r_1 \cdots r_n} \mathrm{X}_1^{r_1} \cdots \mathrm{X}_n^{r_n}, \quad g = \sum b_{r_1 \cdots r_n} \mathrm{X}_1^{r_1} \cdots \mathrm{X}_n^{r_n},$$

则有

$$f + g = h = \sum c_{r_1 \cdots r_n} \mathrm{X}_1^{r_1} \cdots \mathrm{X}_n^{r_n},$$

其中的系数为

$$c_{r_1\cdots r_n} = a_{r_1\cdots r_n} + b_{r_1\cdots r_n},$$

因此

$$\begin{aligned}h(u_1,\cdots,u_n) &= \sum c_{r_1\cdots r_n} u_1^{r_1}\cdots u_n^{r_n} = \sum a_{r_1\cdots r_n} u_1^{r_1}\cdots u_n^{r_n} + \sum b_{r_1\cdots r_n} u_1^{r_1}\cdots u_n^{r_n}\\ &= f(u_1,\cdots,u_n) + g(u_1,\cdots,u_n),\end{aligned}$$

这就证明了 (5). 关系 (6) 用类似方法证明, 不过稍嫌复杂.

此外直接看到性质 (3),(4),(5) 和 (6) 刻画了映射 v 的特征, 并且显然多项式环 $\mathrm{K}[\mathrm{X}_1,\cdots,\mathrm{X}_n]$ 在 v 下的像正是由 K 和 L 的元素 $u_1,\cdots,u_n$ 生成的 L 的子环 $\mathrm{K}[u_1,\cdots,u_n]$.



注 1 同态 v 的核由使得 $f(u_1,\cdots,u_n)=0$ 的多项式组成, 这样的多项式就是由 K 的使得

$$\sum c_{r_1\cdots r_n} u_1^{r_1}\cdots u_n^{r_n} = 0$$

的几乎全部为零的元素的族 $(a_{r_1\cdots r_n})$. 换句话说, v 的核由 §26 第 2 小节所定义的系数在 K 的 $u_1,\cdots,u_n$ 之间的代数关系组成. 在实际中, 对于这些代数关系和系数在 K 内的使得 $f(u_1,\cdots,u_n)=0$ 的多项式不做任何区别.

前面所说的事实表明, 如果 $u_1,\cdots,u_n$ 是 K 上代数无关的, v 的核缩减为 $\{0\}$, 因此 v 是从环 $\mathrm{K}[\mathrm{X}_1,\cdots,\mathrm{X}_n]$ 到 L 的子环 $\mathrm{K}[u_1,\cdots,u_n]$ 的一个同构.

公式 (5) 和 (6) 还可以如下解释. 对于每个多项式 $f\in\mathrm{K}[\mathrm{X}_1,\cdots,\mathrm{X}_n]$, 考虑对应的多项式映射

$$f^*:\mathrm{L}^n\to\mathrm{L},$$

其定义是, 对于任意 $u_i\in\mathrm{L}$,

$$f^*(u_1,\cdots,u_n) = f(u_1,\cdots,u_n).$$

关系 (5) 和 (6) 则写成

$$\begin{aligned}(f+g)^* &= f^*+g^*,\\ (fg)^* &= f^*g^*,\end{aligned}$$

并且表明系数在 K 内的 L^n 上的两个多项式函数的和与乘积仍然是系数在 K 内的多项式函数. 关系 (4) 表明哪些多项式函数是关于 L^n 的典范基的坐标函数, 而 (3) 表明哪些函数是从 L^n 到 K 内的常映射.

事实上, 如果用 M 表示从 L^n 到 L 内的所有映射的环, 那么显然所考虑的多项式映射的集合就是由从 L^n 到 K 内的坐标函数和常映射 (一般把这种映射等同于所对应的 K 的元素本身) 生成的 M 的子环.

3. 无限域的情形

设 K 是一个交换环, f 和 g 是两个系数在 K 内的 n 个未定元的多项式, 而

$$f^*, g^* : \mathrm{K}^n \to \mathrm{K}$$

是 K^n 上对应的多项式函数. 有可能 f^* 和 g^* 是重合的, 而 f 和 g 却是不同的 (这恰恰是一般不能等同系数在 K 内的多项式和 K^n 上的多项式函数的理由).

例 3 取 $n = 1$, 而 K 是 q 个元素的有限域. 由于乘法群 K^* 有 $q-1$ 个元素, §7 的定理 5 表明我们有

$$x^{q-1} = 1 \quad 对于所有\ x \in \mathrm{K}^*,$$

所以

$$x^q = x, \quad 对于所有\ x \in \mathrm{K}.$$

多项式 X^q 和 X 显然是不同的, 却定义 K 上相同的函数.

不过这个困难在经典情形 ($\mathrm{K} = \mathbf{Q}, \mathbf{R}$ 或 $\mathbf{C}$) 不会发生, 因为有下列结果:

定理 1 *设 f 和 g 是两个系数在一个无限整环 K 内的 n 个未定元的多项式. $f = g$, 必须并且只需*

$$f(x) = g(x) \quad 对于所有\ x \in \mathrm{K}^n.$$

考虑多项式 $f - g$, 问题显然归结为指出对于所有多项式 $h \in \mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$, 关系

$$h(x) = 0 \quad 对于所有\ x \in \mathrm{K}^n$$

蕴含 $h = 0$. 我们分几步证明.

引理 1 *设 f 是系数在一个交换环 K 内的一个未定元的多项式. K 的一个元素 a 是 f 的一个根, 必须并且只需存在一个多项式 $g \in \mathrm{K}[\mathrm{X}]$, 使得*

$$f(\mathrm{X}) = (\mathrm{X} - a)g(\mathrm{X}).$$

(这个关系表明在环 K[X] 内 f 是 $\mathrm{X} - a$ 的一个倍式.)

事实上, 设 Y 是异于 X 的一个未定元, 并且用多项式 $a + \mathrm{Y}$ 代换 X, 我们得到 Y 的一个多项式 $f(a + \mathrm{Y})$, 它可以写成

$$f(a + \mathrm{Y}) = u_0 + u_1\mathrm{Y} + \cdots + u_m\mathrm{Y}^m,$$

其中 $u_i \in \mathrm{K}$. 在这个关系中用 0 代换 Y, 显然得到

$$f(a) = u_0,$$

因此对于 Y 的某个多项式 h 可以写出

$$f(a+\mathrm{Y}) = f(a) + \mathrm{Y}\cdot h(\mathrm{Y}).$$

在得到的结果里用 $\mathrm{X}-a$ 代换 Y, 并设 $h(\mathrm{X}-a) = g(\mathrm{X})$, 即得

$$f(\mathrm{X}) = f(a) + (\mathrm{X}-a)\cdot g(\mathrm{X}).$$

如果 $f(a)=0$, 即 a 是 f 的一个根, 此式即表明 $f(\mathrm{X}) = (\mathrm{X}-a)g(\mathrm{X})$; 逆命题是显然的.

引理 2 *设 f 是系数在一个交换环 K 内的一个未定元的次数 $n \geqslant 0$ 的多项式. 如果 K 是整环, 则 f 在 K 内至多有 n 个根.*

如果 f 的次数是 0, 那么 f 是一个非零常量, 不具有任何根, 因此引理当 $n=0$ 时成立. 现在对于 f 的次数 n 进行归纳推理.

设 $a\in \mathrm{K}$ 是 f 的一个根, 则可以写出 $f(\mathrm{X}) = (\mathrm{X}-a)g(\mathrm{X})$, 由于 K 是整环, (根据 §27, 定理 1) 其中 g 是 $n-1$ 次的. 如果 b 是 f 在 K 内的另一个根, 应当有 $0=(b-a)g(b)$, 由于 K 是整环, 我们得到结论: f 在 K 内异于 a 的根都是 g 根. 由于 g 是 $n-1$ 次的, 根据归纳假设它至多有 $n-1$ 根, 因此 f 在 K 内至多有 n 个根, 这就证明了引理.

引理 2 显然在一个未定元的情形直接证明了定理 1, 因为它表明如果基础环 K 是整环, 一个多项式 $h=\mathrm{K}[\mathrm{X}]$ 不能有无穷个根, 除非它是零.

剩下的事情是证明多项式 h 是 n 个未定元的情形, 我们采用关于 n 的归纳推理. 可以写出

$$h(\mathrm{X}_1,\cdots,\mathrm{X}_n) = \sum h_r(\mathrm{X}_1,\cdots,\mathrm{X}_n)\mathrm{X}_n^r,$$

其中的多项式 $h_r\in \mathrm{K}[\mathrm{X}_1,\cdots,\mathrm{X}_{n-1}]$. 假定对于所有 $x\in\mathrm{K}^n$ 有 $h(x)=0$, 那么显然对于所有 $y\in\mathrm{K}^{n-1}$ 和所有 $t\in\mathrm{K}$ 就有

$$\sum h_r(y)t^r = 0.$$

对于给定的 $y\in\mathrm{K}^{n-1}$, 我们发现一个未定元的多项式

$$\sum h_r(y)\mathrm{T}^r$$

在 K 的所有点是零, 由于定理 1 已经对于一个未定元的多项式证明, 由此推出

$$h_r(y) = 0 \quad \text{对于所有 } y\in\mathrm{K}^{n-1} \text{ 和所有 } r.$$

根据归纳假设对于所有 r 有 $h_r=0$, 最终有关系 $h=0$, 这就结束了证明.

事实上, 可以改进定理 1, 参见习题 1.

§§27, 28 习题

1. 设 K 是一个无限整环. 称 K^n 的一个子集 A 是在 K^n 内的 **Zariski 开集**, 如果存在有限个多项式 $f_1, \cdots, f_r \in K[X_1, \cdots, X_n]$, 使得 A 在 K^n 内的补集是满足

$$f_1(x) = \cdots = f_r(x) = 0$$

的 $x \in K^n$ 集合. 有了这个定义后, 设 f 和 g 是两个系数在 K 内的 n 个未定元的多项式; 假定在 K^n 内存在一个非空 Zariski 开集 A, 使得

$$f(x) = g(x) \quad \text{对于所有 } x \in A,$$

证明 $f = g$ (代数恒等式延拓原则).

2. 证明如果三个多项式 $f, g, h \in \mathbf{R}[X]$ 满足下列方程中的任何一个, 则有 $f = g = h = 0$:

$$f(X)^2 - Xg(X)^2 = Xh(X)^2,$$
$$f(X)^2 - Xg(X)^2 + h(X)^2 = 0,$$
$$f(X)^2 + g(X)^2 + (X+2)h(X)^2 = 0.$$

前面的 $\mathbf{R}$ 换成任何一个交换域可以吗?

3. 设 K 是一个交换域, f 是系数在 K 内的一个未定元的多项式, 而 $a_1, \cdots, a_r$ 是 f 在 K 内的两两不同的根. 借助 §28 的引理 1 证明存在一个系数在 K 内的 g 多项式, 使得

$$f(X) = (X - a_1) \cdots (X - a_r) g(X).$$

应用: 计算 (实际用不着计算!) 行列式

$$\begin{vmatrix} 1 & 1 & 2 & 3 \\ 1 & 2-x^2 & 2 & 3 \\ 2 & 3 & 1 & 5 \\ 2 & 3 & 1 & 9-x^2 \end{vmatrix}.$$

4. 对于下列行列式回答同样问题:

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1-x & 1 & \cdots & 1 \\ 1 & 1 & 2-x & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & n-x \end{vmatrix}.$$

5. 设 $f_1, \cdots, f_n$ 是其系数在交换环 K 内的一个未定元的多项式, 它们的次数至多是 $n-2$. 证明对于任意 $x_i \in K$ 有

$$\begin{vmatrix} f_1(x_1) & f_1(x_2) & \cdots & f_1(x_n) \\ \vdots & \vdots & & \vdots \\ f_n(x_1) & f_n(x_2) & \cdots & f_n(x_n) \end{vmatrix} = 0.$$

¶6. 设 K 是一个无限交换域, 而 $a_0, \cdots, a_n$ 是 K 的两两不同的元素. 证明存在唯一的一个至多 n 次的多项式 $f \in \mathrm{K}[\mathrm{X}]$, 满足

$$f(a_i) = b_i \quad (0 \leqslant i \leqslant n).$$

其中的 b_i 是 K 的给定元素, 而 f 由 **Lagrange 插值公式**提供:

$$f(\mathrm{X}) = \sum_{i=0}^{n} b_i \frac{(\mathrm{X} - a_0) \cdots (\mathrm{X} - a_{i-1})(\mathrm{X} - a_{i+1}) \cdots (\mathrm{X} - a_n)}{(a_i - a_0) \cdots (a_i - a_{i-1})(a_i - a_{i+1}) \cdots (a_i - a_n)}.$$

例子: 求一个三次多项式, 使得

$$f(1) = 2, \quad f(2) = 1, \quad f(3) = 4, \quad f(4) = 3.$$

¶7. 令

$$z_k = \cos \frac{2k\pi}{n} + i \cdot \sin \frac{2k\pi}{n},$$

求一个复系数的 $n-1$ 次多项式, 使得

$$f(z_k) = k + 1 \quad \text{对于 } 0 \leqslant k \leqslant n-1.$$

$$\left[\text{答案: } f(\mathrm{X}) = \frac{n+1}{2} - \frac{1}{2} \sum_{k=1}^{n-1} \left(1 - i \cdot \cot \frac{k\pi}{n}\right) \mathrm{X}^k.\right]$$

8. 设 f 是定义在自然数集合 $\mathbf{N}$ 上的复数值的一个函数. 用

$$\Delta f(n) = f(n+1) - f(n)$$

定义一个新的函数 Δf. 逐次定义

$$\Delta^2 f = \Delta(\Delta f), \quad \Delta^3 f = \Delta(\Delta^2 f), \quad \cdots$$

最后, 称 f 是 r **次多项式函数**, 如果存在常数 $a_0, \cdots, a_r$, 使得

$$f(n) = a_r n^r + a_{r-1} n^{r-1} + \cdots + a_0 \quad \text{对于所有 } n \in \mathbf{N},$$

其中的 $a_r \neq 0$.

a) 证明如果 f 是 r 次多项式函数, 则

$$\Delta^{r+1} f = 0, \quad \Delta^r f \neq 0.$$

b) 对于以下的 f 计算 Δf:

$$f(n) = \frac{n(n-1) \cdots (n-r+1)}{r!} = \binom{n}{r} \quad \text{对于 } n \in \mathbf{N};$$

由此推出如果 f 是 r 次的任意多项式函数, 则有

$$f(n) = c_0 + c_1 \binom{n}{1} + \cdots + c_r \binom{n}{r}, \quad \text{其中 } c_k = \Delta^k f(0).$$

c) 证明: $\mathbf{N}$ 上的一个多项式函数 f 如果满足条件

$$\Delta^{r+1} f = 0, \quad \Delta^r f \neq 0,$$

则 f 是一个 r 次多项式函数.

d) 设 g 是 $\mathbf{N}$ 上的一个 r 次多项式函数. 证明存在唯一的一个 $\mathbf{N}$ 上的 r 次多项式函数, 使得

$$\Delta f = g, \quad f(0) = 0.$$

通过用问题 b) 的公式计算 f, 由此推出和

$$g(0) + g(1) + \cdots + g(n)$$

的一个表达式.

应用: 计算和

$$1^2 + 2^2 + \cdots + n^2, 1^3 + 2^3 + \cdots + n^3.$$

¶¶9. (我们回忆几个定义. 如果 A 是一个交换环, A 的一个理想 I 是*素理想*, 如果 $\mathrm{I} \neq \mathrm{A}$, 并且对于 $x, y \in \mathrm{A}$, 关系 $xy \in \mathrm{I}$ 蕴含 $x \in \mathrm{I}$ 或 $y \in \mathrm{I}$; 称 A 的一个理想 I 是*极大的*, 如果 $\mathrm{I} \neq \mathrm{A}$, 并且包含 I 的 A 的仅有的理想是 I 和 A; 最后 A 的不同于 A 的所有理想都至少包含在一个极大的理想内). 我们打算证明, *如果 K 是一个交换环, 则 K 的所有素理想的交集是 K 的幂零元的集合.*

a) 证明, 如果 K 的一个理想 I 满足 $\mathrm{I} \neq \mathrm{K}$, 那么由 I 生成的多项式环 K[X] 的理想 I′ 满足 $\mathrm{I}' \neq \mathrm{K[K]}$. 由此推出对于 K 的所有素理想, 存在包含它的 K[X] 的极大理想.

b) 假定 $u \in \mathrm{K}$ 属于 K 的所有素理想. 证明多项式 $1 - u\mathrm{X}$ 不属于环 K[X] 的任何极大理想. 由此推出它在环 K[X] 内是可逆的.

c) 证明, 多项式 $1 - u\mathrm{X}(u \in \mathrm{K})$ 在环 K[X] 内是可逆的当且仅当 u 是幂零的. 由此推出所宣布的定理.

d) 设 I 是交换环 K 的一个理想, 并且 $\mathrm{I} \neq \mathrm{K}$. 证明包含 I 的 K 的素理想的交集由这样的 $x \in \mathrm{K}$ 组成, 对于至少一个整数 n 有

$$x^n \in \mathrm{I}.$$

10. 设 K 是一个交换环. 由一个多项式 $f \in \mathrm{K[X]}$ 生成的 K[X] 的子环 K[f] 是整个 K[X], 必须并且只需

$$f(\mathrm{X}) = a\mathrm{X} + b, \quad a \neq 0.$$

¶11. 设 K 是一个交换环. 称 K 的元素的所有序列

$$f = (a_0, a_1, \cdots, a_n, \cdots)$$

为**系数在 K 内的一个未定元的形式幂级数** (没有假设 a_i 几乎全部为零). 借助定义多项式的和与积的 §27 第 2 小节的公式 (2) 和 (3) 定义两个这样的形式幂级数的和与积. 证明采用这样的定义就得到了一个*交换环*, 它包含一个同构于 K[X] 的子环.

这样得到的环通常记作 K[[X]]; 代替初始的记号 $f = (a_0, a_1, \cdots, a_n, \cdots)$, 采用写法

$$f = a_0 + a_1\mathrm{X} + \cdots + a_n\mathrm{X}^n + \cdots = \sum_{n=0}^{\infty} a_n\mathrm{X}^n, \tag{$*$}$$

这种写法便于更容易地记住定义基本运算的公式. 为了求两个幂级数的乘积, 把它们 "逐项" 相乘, 再在所得到的结果里把次数相同的项合并在一起. 当然公式 $(*)$ 没有任何理论上的意义, 因为它可能含有无穷多个非零项, 它只不过是表现 $a_i \in \mathrm{K}$ 的级数的一个方便的记号.

证明下列结果:

a) 环 $\mathrm{K}[[\mathrm{X}]]$ 是整环, 必须并且只需 K 是整环.

b) $\mathrm{K}[[\mathrm{X}]]$ 的一个元素 $(*)$ 在 $\mathrm{K}[[\mathrm{X}]]$ 是可逆的, 必须而且只需它的 "常项" a_0 在 K 内是可逆的 (这是多项式环和形式幂级数环之间的主要差别).

c) 计算 $1-\mathrm{X}$ 在 $\mathrm{K}[[\mathrm{X}]]$ 内的逆.

¶¶ 12. 设 K 是一个交换环, 而

$$p(\mathrm{X},\mathrm{Y}) = p_0(\mathrm{X}) + p_1(\mathrm{X})\mathrm{Y} + \cdots + p_n(\mathrm{X})\mathrm{Y}^n$$

是一个系数在 K 内的两个未定元的多项式. 假定 $p_0(0)=0$ 和 $p_1(0)$ 在 K 内是可逆的 [如果 K 是一个域, $p(0,0)=0$, 并且 $p'_{\mathrm{Y}}(0,0)\neq 0$]. 证明存在唯一的一个系数在内 K 内并且没有常项的形式幂级数

$$y = a_1\mathrm{X} + a_2\mathrm{X}^2 + \cdots$$

满足关系 $p(\mathrm{X},y)=0$.

[可以如下进行: 假定求出常量 $a_1,\cdots,a_r \in \mathrm{K}$, 使得多项式

$$p(\mathrm{X}, a_1\mathrm{X} + a_2\mathrm{X}^2 + \cdots + a_r\mathrm{X}^r)$$

不含有次数 $\leqslant r$ 的项, 证明存在 $a_{r+1} \in \mathrm{K}$, 使得

$$p(\mathrm{X}, a_1\mathrm{X} + a_2\mathrm{X}^2 + \cdots + a_r\mathrm{X}^r + a_{r+1}\mathrm{X}^{r+1})$$

不含有次数 $\leqslant r+1$ 的项.]

如果 $\mathrm{K} = \mathbf{C}$ 和 $f(\mathrm{X},\mathrm{Y}) = (\mathrm{X}-1)^p - \mathrm{Y}^q$, 其中的 p 和 q 是正整数. 你发现这样得到的形式幂级数和在分析里的函数

$$(z-1)^{p/q}$$

的幂级数展开的关系了吗?

¶ 13. 设 p 是一个素数, f 和 g 是系数为有理整数的多项式.

a) 证明, 如果 p 整除 fg 的所有系数, 则它或整除 f 的所有系数, 或整除 g 的所有系数.

b) 说系数为有理整数的多项式是**本原多项式**, 如果它的系数的最大公约数等于 1. 证明如果 f, g 是本原多项式, 则它们的乘积同样是本原的.

c) (Gauss 引理) 给定一个系数为有理整数的多项式 h, 称其系数的最大公约数为 h 的**容度**, 记作 $c(h)$. 证明

$$c(fg) = c(f)c(g) \quad \text{对于任意 } f, g \in \mathbf{Z}[\mathrm{X}].$$

¶ 14. 设 K 是一个交换环, $\mathfrak{p}$ 是 K 的一个理想, 而 f, g 是系数在 K 内的两个多项式. 假定 fg 的系数在 $\mathfrak{p}$ 内. 证明 $\mathfrak{p}$ 含有 f 的所有系数, 或 g 的所有系数.

¶ 15. 设 K 是一个交换整环.

a) 设 f 和 g 是其系数在 K 内的两个非常量多项式. 在系数在 K 内的两个未定元的多项式环 $\mathrm{K}[\mathrm{X}, \mathrm{Y}]$ 内, 考虑由多项式 $f(\mathrm{X})$ 和 $g(\mathrm{Y})$ 生成的理想 I. 证明

$$\mathrm{I} \neq \mathrm{K}[\mathrm{X}, \mathrm{Y}].$$

[假定有一个形式为

$$u(\mathrm{X}, \mathrm{Y}) f(\mathrm{X}) + v(\mathrm{X}, \mathrm{Y}) g(\mathrm{Y}) = 1$$

的关系, 并且考查左端的最高次的齐次项.]

b) 设 $f_1, \cdots, f_n$ 是系数在 K 内的一个未定元的多项式. 证明由 $f_1(\mathrm{X}_1), \cdots, f_n(\mathrm{X}_n)$ 生成的环 $\mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$ 的理想不是整个环 $\mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$, 如果 f_i 都不是常元.

c) 证明对于所有满足条件 $1 \leqslant k \leqslant n$ 的整数 k, 由 $\mathrm{X}_1, \cdots, \mathrm{X}_k$ 生成的 $\mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$ 的理想是素理想. 这 n 个理想是两两不同的吗?

¶16. 设 K 是一个交换环. 证明 K 的下列性质是等价的:

(i) K 没有非零的幂零元;

(ii) K[X] 的所有可逆元是常元 (参见本节的习题 9 和 §8 的习题 1; 还可以考虑和本节的习题 11 的关系).

17. 设 V 和 W 是交换域 K 上的两个向量空间. 说从 V 到 W 内的一个映射是**多项式的**, 如果存在 V 的一个基和 W 的一个基, 使得向量 $y = f(x) \in \mathrm{W}$ 的坐标由形如

$$\eta_j = p_j(\xi_1, \cdots, \xi_m) \quad (1 \leqslant j \leqslant n)$$

的公式由向量 $x \in \mathrm{V}$ 的坐标给定, 其中 p_j 是其系数在 K 内的 $m = \dim(\mathrm{V})$ 个未定元的多项式. 此外称 f 是 **r 次齐次的**, 如果各个 p_j 是 r 次齐次的.

证明这些定义不依赖 V 和 W 的基的选取 (即如果所宣布的条件对于这些基的一个特殊选取是满足的, 那么它们对于所有其他的选取也是满足的). 证明, 如果域 K 是有限的, 则所有从 V 到 W 内的映射都是多项式的 (在这种情形这个概念就失去许多益处). 以下假定 K 是*无限的*.

用 $\mathrm{S}(\mathrm{V}, \mathrm{W})$ 表示从 V 到 W 内的多项式映射的集合, 而用 $\mathrm{S}_r(\mathrm{V}, \mathrm{W})$ 表示 r 次齐次多项式映射的集合. 最后令

$$\mathrm{S}(\mathrm{V}) = \mathrm{S}(\mathrm{V}, \mathrm{K}), \quad \mathrm{S}_r(\mathrm{V}) = \mathrm{S}_r(\mathrm{V}, \mathrm{K}),$$

$\mathrm{S}(\mathrm{V})$ (对应的, $\mathrm{S}_r(\mathrm{V})$) 的元素称为 V 上的**多项式** (对应的, **r 次齐次多项式**) **函数**.

证明所有的 $f \in \mathrm{S}(\mathrm{V}, \mathrm{W})$ 按唯一的方式写成形式

$$f = f_0 + f_1 + \cdots,$$

其中的 f_r 是 r 次齐次多项式, 对于几乎所有的 r, $f_r = 0$.

证明 $\mathrm{S}(\mathrm{V})$ 是从集合 V 到域 K 的所有映射的环的一个子环, 并且 $\mathrm{S}(\mathrm{V})$ 含有线性映射和常映射 (习惯上把它们等同于 K 的元素, 以至于 K 典范地等同于环 $\mathrm{S}(\mathrm{V})$ 的一个子域). 设 $f_1, \cdots, f_m$ 是关于 V 的一个基 V 的坐标函数, 证明

$$\mathrm{S}(\mathrm{V}) = \mathrm{K}[f_1, \cdots, f_m].$$

并且元素 $f_1, \cdots, f_m$ 在 K 上是代数无关的.

证明 S(V, W) 是从集合 V 到向量空间 W 的所有映射的向量空间的一个向量子空间. 证明, 如果 $f \in \mathrm{S(V)}$ 并且 $g \in \mathrm{S(V, W)}$, 则从 V 到 W 由

$$h(x) = f(x)g(x) \quad 对于所有\ x \in \mathrm{V}$$

定义的映射 $h = fg$ 还是多项式的. 由此推出可以把 S(V, W) 看作环 S(V) 上的一个模. 设 (a_i) 是 V 的一个基, (b_j) 是 W 的一个基, 用 f_{ij} 表示从 V 到 W 的线性映射, 它满足条件

$$f_{ij}(a_k) = \begin{cases} b_j, & 如果\ k = i, \\ 0, & 如果\ k \neq i; \end{cases}$$

证明 f_{ij} 组成 S(V) 上的模 S(V,W) 的一个基.

设 U, V, W 是 K 上的有限维向量空间, 而

$$f : \mathrm{U} \to \mathrm{V}, \quad g : \mathrm{V} \to \mathrm{W}$$

是两个多项式映射. 证明复合映射 $g \circ f$ 是多项式的. 如果 f 和 g 分别是 r 次和 s 次齐次的, 则 $g \circ f$ 是 rs 次齐次的.

18. 设 K 是一个无限交换域. 考虑从 K 到 K^3 内的由

$$f(t) = (t^2 + t + 1, t^3 + t + 1, t^4 + t + 1)$$

给定的多项式映射. 求 K^3 上的在 $f(\mathrm{K})$ 上的所有的点取零值的所有多项式函数. 使得这些函数是零的 K^3 的点是什么?

对于从 K^* 到 K^3 的由

$$f(t) = \left(\frac{t+1}{t}, \frac{t^2+1}{t}, \frac{t^3+1}{t}\right)$$

定义的映射回答同样的问题.

¶19. 设 K 是一个交换环, 而 M 是 K 上的模. 我们打算把 M"嵌入" 到系数在 K 内的一个未定元的多项式环 K[X] 内. 为此, 考虑记作 M[X] 的集合, 其元素是 M 的几乎全部为零的元素的序列

$$(m_0, m_1, \cdots).$$

用公式

$$(m_0', m_1', \cdots) + (m_0'', m_1'', \cdots) = (m_0' + m_0'', m_1' + m_1'', \cdots)$$

定义 M[X] 的加法, 最后令

$$(a_0, a_1, a_2, \cdots) \cdot (m_0, m_1, m_2, \cdots) = (a_0 m_0, a_0 m_1 + a_1 m_0, \cdots)$$

定义 M[X] 的元素乘以 K[X] 的元素的乘积. 在这样的定义之下证明集合 M[X] 实际上是一个 K[X]-模, 称为**系数在 M 内的一个未定元的多项式模**. 把每一个 $m \in \mathrm{M}$ 等同于 M[X] 的元素 $(m, 0, \cdots)$, 证明在 K[X]-模 M[X] 内有

$$(m_0, m_1, m_2, \cdots) = m_0 + m_1 \mathrm{X} + m_2 \mathrm{X}^2 + \cdots$$

(注意: 这里把标量, 即 K[X] 的元素写在了 M[X] 的元素的右边, 以便符合在多项式内系数写在单项式左边的传统).

设 M 和 N 是两个 K-模, 而 u 是从 M 到 N 内的一个同态, 证明存唯一的 K[X]-模的同态

$$\bar{u}: \mathrm{M[X]} \to \mathrm{N[X]},$$

在 M 上它跟 u 重合.

假定 M 是有限生成自由的; 证明 M[X] 是一个有限生成自由的 K[X]-模, 并且 M 在 K 上的基也是 M[X] 在 K[X] 上的基 (关于这个构造的应用, 参见 §35, 习题 10).

¶20. 设 K 是一个交换环, E 是 K 上的一个模, 而 u 是 E 的一个自同态. 给定一个系数在 K 内的多项式

$$f(\mathrm{X}) = a_0 + a_1\mathrm{X} + \cdots + a_n\mathrm{X}^n,$$

令

$$f(u) = a_0 \cdot j_{\mathrm{E}} + a_1 u + \cdots + a_n u^n,$$

由此得到 E 的一个新的自同态. 现在考虑从乘积集合 $\mathrm{K[X]} \times \mathrm{E}$ 到集合 E 内的由

$$(f, x) = f(u)(x)$$

给定的映射. 证明配备了运算 $(x, y) \to x + y$ 和刚定义的映射, 集合 E 是环 K[X] 上的一个模, 记为 E_u.

反之, 设 M 是一个 K[X]-模, 再设 E 是通过限制标量(*)到环 K 上得到的 M 的缩减 K-模. 考虑 M 内的比例为 X 的位似为从 E 到 E 内的映射 u. 证明 u 是 K-模 E 的一个自同态, 并且 $\mathrm{M} = \mathrm{E}_u$. 换句话说, 环 K[X] 上的一个模等同于一个序偶, 它由环 K 上的一个模和这个 K-模的一个自同态组成. (这个结果表明 K-模的自同态的研究可以归结为 K[X]-模的研究, 在 §35 的习题内将会确信这一看法.)

给定 E 和 u 如前, K[X]-模 E_u 的子模是什么?

考虑两个 K-模 E 和 F, E 的一个自同态 u, 和 F 的一个自同态 v, 从 K[X]-模 E_u 到 K[X]-模 E_v 的同态是什么?

¶21. 模仿上一个习题的构造, 证明如果 K 是一个交换环, 在多项式环 K[X, Y] 上的模等同于一个三元组 (E, u, v), 它由一个 K-模 E 和 E 的两个自同态 u 和 v 组成, 这里 u 和 v 满足关系 $u \circ v = v \circ u$. 可以推广到 n 个未定元的情形吗?

¶¶22. 设 L 是一个交换域, 而 K 是 L 的一个子域.

a) 考虑系数 a_{ij} 和右端 b_i 在 K 内的一个线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1, \\ \qquad\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m, \end{cases} \tag{$*$}$$

和 n 个未定元的系数在 K 内的多项式 p. 假定存在方程组 $(*)$ 的一个解

$$(x_1, \cdots, x_n) \in \mathrm{L}^n,$$

(*) 设 L 是一个环, K 是 L 的一个子环, 而 M 是一个左 L-模. **限制标量到** K 上的 M 的缩减模这样得到: 考虑加法群 M 和从 $\mathrm{K} \times \mathrm{M}$ 到 M 内的映射 $(a, m) \to am$, 它是定义 M 的 L-模结构的从 $\mathrm{L} \times \mathrm{M}$ 到 M 内的映射到 $\mathrm{K} \times \mathrm{M}$ 上的限制. 换句话说, 保持 "向量" 和加法的原样, 但是仅保留比例属于 K 的位似变换. 例如, 所有的复向量空间也定义一个实向量空间.

它使得 $p(x_1,\cdots,x_n)\neq 0$. 证明, 如果 K 是无限域, 则在 K^n 内存在 (*) 的一个解使得 p 不为零. (以 $\mathrm{K}=\mathbf{R}$ 和 $\mathrm{L}=\mathbf{C}$ 为例从几何上说明这个结果.)

b) 考虑两个矩阵 $\mathrm{A},\mathrm{B}\in\mathrm{M}_n(\mathrm{K})$. 假定存在一个矩阵 $\mathrm{V}\in\mathrm{GL}(n,\mathrm{L})$, 使得 $\mathrm{B}=\mathrm{VAV}^{-1}$. 证明存在 $\mathrm{U}\in\mathrm{GL}(n,\mathrm{K})$, 使得 $\mathrm{B}=\mathrm{UAU}^{-1}$ (考虑方程 UA = BU, 其中 $\det(\mathrm{U})\neq 0$, 并且应用前一个问题). [注意, 这个结果当 K 是一个有限域时仍然成立, 但是证明明显地更加困难; 参见 §35, 习题 12.]

¶23. 设 K 是一个交换环, 而 I 是由 X^{n+1} 生成的 K[X] 的理想. 描述商环 $\mathrm{L}=\mathrm{K}[\mathrm{X}]/\mathrm{I}$ (可以证明 L 是由 K 和一个满足关系 $\varepsilon^{n+1}=0$ 的元素 ε 生成的).

求它的可逆元. [当 $\mathrm{K}=\mathbf{C}$ 时, 环 L 涉及 n 阶有限展开理论.]

§29 有理分式

1. 整环的分式域: 预备知识

显然, 一个域的所有子环都是整环. 反之我们要问, 是否所有整环都是一个域的子环? 在本节我们要阐明, 至少在整环是交换的情形下, 可以给这个问题一个肯定的回答.

为了构建包含给定交换整环 K 的一个域, 我们先假定问题已经解决, 即假定构建了一个域 L, K 是它的一个子环. 那么 K 的所有非零元 a 在 L 内拥有一个逆元. 更一般地说来, 给定两个元素 $a,b\in\mathrm{K}$, 其中的 $b\neq 0$, 在 L 里可以考虑分式

$$a/b=ab^{-1}.$$

这些分式的集合是 L 的包含 K 的一个子域. 事实上, 设 L′ 是这个集合, 考虑 L′ 的两个元素 x,y, 可以把它们表示为

$$x=ab^{-1},\quad y=cd^{-1},$$

其中的 a,b,c,d 是 K 的元素, 并且 $b\neq 0,d\neq 0$. 一个平凡的计算 (考虑到 K 的交换性) 表明有关系

$$x+y=(ad+bc)/bd,\tag{1}$$

$$yx=xy=ac/bd,\tag{2}$$

并且对于所有 $a\in\mathrm{K}$ 显然有

$$a/1=a.\tag{3}$$

这样我们已经发现 L′ 是 L 的包含 K 的一个子环. 为了确认 L′ 是 L 的一个子域, 我们注意当且仅当 $a\neq 0$ 时, L′ 的一个元素 $x=a/b$ 不是零. 于是有

$$x^{-1}=(ab^{-1})^{-1}=(b^{-1})^{-1}a^{-1}=ba^{-1}=b/a,$$

显然这就证实了 x 在 L' 内是可逆的.

如果可能 "嵌入" K 到一个域 L 内, 可以假定 (需要时利用上述的构造, 用 L' 代替 L) L 的所有元素可以写成形式

$$x = a/b, \quad \text{其中 } a, b \in \mathrm{K}, \quad b \neq 0. \tag{4}$$

用 F 表示序偶 (a, b) 的集合, 其中的 $a, b \in \mathrm{K}, b \neq 0$, 那么由

$$\nu(a, b) = a/b = ab^{-1}$$

定义的映射 $\nu : \mathrm{F} \to \mathrm{L}$ 就是满射的. 用 R 表示关联于映射 ν (§4, 第 1 小节) 的集合 F 上的等价关系, 即关系

$$\nu(a, b) = \nu(c, d),$$

此式还可以写成 $a/b = c/d$, 即

$$ad = bc,$$

那么 ν 就是由从 F 到 F/R 上的典范映射和从 F/R 到 L 上的一个映射 ν' 复合而成 (§4, 定理 2), 并且显然 ν' 是双射的. 如果我们利用 ν' 把 L 的元素等同于它们对应的 F/R 的元素, 那么就会发现可以把 F/R 看作一个域, 其中的运算由 (1) 和 (2) 定义. 更精确地说, 如果 F/R 的元素 x 和 y 由 F 内的 (a, b) 和 (c, d) 代表, 那么 $x + y$ 将是 F/R 的由序偶 $(ad + bc, bd)$ 代表的元素, 而 xy 将是 F/R 的由序偶 (ac, bd) 代表的元素.

这些考虑, 当然是假定了问题已经解决, 将引导我们在下一小节从环 K 出发构建包含 K 的一个域 L.

2. 分式域的构造

设 K 是一个交换整环. 用 F 表示序偶 (a, b) 的集合, 其中 $a, b \in \mathrm{K}, b \neq 0$. 给定 F 的元素 (a, b) 和 (c, d), 用记号

$$(a, b) \equiv (c, d) \mod \mathrm{R}$$

表示关系

$$ad = bc.$$

先来阐明 R 是 F 上的一个等价关系.

首先, 关系

$$(a, b) \equiv (a, b) \mod \mathrm{R}$$

总是成立的, 这是因为它就是

$$ab = ba.$$

其次, 关系

$$(a,b)\equiv(c,d)\quad \bmod \mathrm{R}$$

蕴含关系

$$(c,d)\equiv(a,b)\quad \bmod \mathrm{R}.$$

这是因为第一个写成 $ad=bc$, 而第二个写成 $cb=da$.

最后考虑关系

$$(a,b)\equiv(c,d)\quad \bmod \mathrm{R},$$
$$(c,d)\equiv(e,f)\quad \bmod \mathrm{R},$$
$$(a,b)\equiv(e,f)\quad \bmod \mathrm{R},$$

它们可以改写为 $ad=bc, cf=de$ 和 $af=be$. 第一个关系乘以 f, 而第二个关系乘以 b (为了使得在所考虑的两个关系里都出现 bcf) 即得 $adf=bde$, 或 $(af-be)d=0$. 由于 $d\neq 0$ 并且 K 是整环, 故 $af-be=0$, 因此前两个关系蕴含第三个关系.

我们已经证明了 R 是 F 上的一个等价关系, 这就可以让我们构造一个商集 F/R, 用 θ 表示从 F 到 F/R 上的典范映射.

现在我们要指出存在集合 F/R 上的两个运算

$$(x,y)\to x+y \quad 和 \quad (x,y)\to xy$$

使得对于任意 $(a,b),(c,d)\in \mathrm{F}$ 有

$$\theta(a,b)+\theta(c,d)=\theta(ad+bc,bd), \tag{5}$$
$$\theta(a,b)\cdot\theta(c,d)=\theta(ac,bd). \tag{6}$$

首先注意关系 (5) 和 (6) 的右端有意义, 即有 $bd\neq 0$: 这来源于 $b\neq, d\neq 0$ 和 K 是一个整环这一事实.

现在为了确认从 $\mathrm{F/R}\times\mathrm{F/R}$ 到 F/R 内的满足条件 (5) 和 (6) 的运算的存在性, 我们要利用 §4 的定理 3, 取 $\mathrm{X}=\mathrm{Y}=\mathrm{Z}=\mathrm{F}, \mathrm{R}=\mathrm{S}=\mathrm{T}$, f 作为映射 $\mathrm{F}\times\mathrm{F}\to\mathrm{F}$, 或由

$$f[(a,b),(c,d)]=(ad+bc,bd),$$

定义或由

$$f[(a,b),(c,d)]=(ac,bd)$$

定义根据所提到的定理, 事情归结为证实下列结果: 关系

$$(a',b')\equiv(a'',b'')\quad \bmod \mathrm{R}$$

和

$$(c',d') \equiv (c'',d'') \mod \mathrm{R}$$

蕴含关系

$$(a'd'+b'c',b'd') \equiv (a''d''+b''c'',b''d'') \mod \mathrm{R} \tag{7}$$

和

$$(a'c',b'd') \equiv (a''c'',b''d'') \mod \mathrm{R}. \tag{8}$$

先证明 (8). 它可以改写为

$$a'c'b''d'' = b'd'a''c'',$$

由于根据假设有

$$a'b'' = b'a'' \quad \text{和} \quad c'd'' = d'c'', \tag{9}$$

要找的关系由刚写出的两个关系边边相乘得到. 至于关系 (7), 它还可以改写为

$$(a'd'+b'c')b''d'' = (a''d''+b''c'')b'd',$$

它明显地等价于关系

$$(a'b''-a''b')d'd''+(c'd''-c''d')b'b''=0,$$

此式显然由 (9) 推出.

于是我们显然处于应用 §4 的定理 3 的条件之下, 从而在商集 F/R 上存在满足条件 (5) 和 (6) 的运算, 并且这些条件像定理所指明的那样无歧义地确定了这两个运算.

注 1 读者不要以为在从整数构造有理数的 "初等" 情形可以简化上述推理. 一个有理数有无数种不同方式写成分数形式, (举例说) 不能局限于令



$$\frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd}$$

定义两个有理数的和, 因为为了使这个公式能够定义两个有理数 (不仅是两个分数的和, 分数本身是完全不感兴趣的概念) 的和, 应当指出当把 a/b 和 c/d 用等价的分数 (即定义同样有理数的分数) 代换时, 右端也用等价的分数代换. 例如, 应当证明分数

$$1/2+4/6=14/12 \quad \text{和} \quad 4/8+2/3=28/24$$

是等价的.

在初等教育阶段, 给出这样的证明是不必要的, 以致这样得到的对两个有理数的和与乘积的定义是没有任何数学价值的, 这在教学上对天真的孩子们是一种很大的欺骗, 掩盖了问题真正的困难.

事实上, 不管采用什么方法都不得不利用 §4 定理 3 的推理, 以及证明 (7) 和 (8) 的极其简单的计算; 一个企图绕过这些过程的构造必定是谬误的.

3. 域的公理的验证

在这一小节我们要证明配备了前一小节所定义的两个运算, 集合 F/R 是一个域.

首先证明加法是交换的. 考虑到 (5), 为此只需注意到如果交换 (a,b) 和 (c,d), 即用 c,d,a,b 代换 a,b,c,d, F 的元素 $(ad+bc,bd)$ 不变.

现在指出加法是结合的. 给定 F/R 的元素

$$x=\theta(a,b),\quad y=\theta(c,d),\quad z=\theta(e,f),$$

我们有

$$(x+y)+z=\theta(ad+bc,bd)+\theta(e,f)=\theta[(ad+bc)f+bde,bdf],$$
$$x+(y+z)=\theta(a,b)+\theta(cf+de,df)=\theta[adf+b(cf+de),bdf].$$

只需验证

$$(ad+bc)f+bde=adf+b(cf+de),$$

而这是容易的.

加法具有一个中性元, 即

$$0=\theta(0,1).$$

事实上, 按照 (5),

$$\theta(0,1)+\theta(a,b)=\theta(0\cdot b+1\cdot a,1\cdot b)=\theta(a,b).$$

F/R 的每个元素

$$x=\theta(a,b)$$

都有一个相反元, 即

$$-x=\theta(-a,b).$$

事实上, 我们有

$$\theta(a,b)+\theta(-a,b)=\theta(0,b^2),$$

因此只需证明

$$\theta(0,b)=\theta(0,1)\quad 对于任意\ b\in \mathrm{K},b\neq 0.$$

这个关系可以改写为 $0\cdot 1=b\cdot 0$, 而这是平凡的结果.

如此一来, 我们已经证明了配备了加法的 F/R 是一个交换群.

现在证明在 F/R 里乘法是交换的, 结合的, 并且具有一个中性元. 根据公式 (6) 交换性是显然的. 观察

$$[\theta(a,b)\theta(c,d)]\theta(e,f) = \theta(ac,db)\theta(e,f) = \theta[(ac)e,(bd)f]$$

和

$$\theta(a,b)[\theta(c,d)\theta(e,f)] = \theta(a,b)\theta(ce,df)\theta = \theta[a(ce),b(df)],$$

即可得到结合性. 最后, 通过平凡的计算即可确信 F/R 的元素

$$1 = \theta(1,1)$$

是乘法的中性元.

为了完成 F/R 是一个域这一事实的证明还需要验证分配律 $(x+y)z = xz+yz$, 以及指出 F/R 的所有非零元是可逆的. 第一点留给读者作为习题. 至于第二点, 首先注意关系

$$\theta(a,b) = 0$$

等价于 $a=0$, 这是因为已经知道 $0=\theta(0,1)$, 所提到的关系还可以写成 $a\cdot 1 = b\cdot 0$. 因此, 如果 x 是 F/R 的非零元, 则有

$$x = \theta(a,b), \quad \text{其中 } a \neq 0, b \neq 0.$$

于是有权考虑 F/R 的元素

$$x^{-1} = \theta(b,a),$$

而这个元素确实是 x 的逆元, 事实上有

$$x \cdot x^{-1} = \theta(ab,ab),$$

故问题归结为指出对于所有 $c \neq 0$ 有

$$\theta(c,c) = 1 = \theta(1,1).$$

而这是显然的, 因为所考虑的关系可以写成 $c\cdot 1 = 1\cdot c$.

4. 环 K 嵌入到它的分式域

为了完全解决在第一小节所提出的问题, 留给我们要做的是指出怎样考虑 K 为域 F/R 的一个子环. 这里跟许多其他情形一样, 我们并不是要真的证明 K 是 F/R 的一个子环 (它甚至不是 F/R 的一个子集), 而是要构造从 K 到 F/R 的一个子环的 "典范" 的同构 (读者回忆一下下列事实将是有益的, 例如把实数等同于复数, 或一个交换环的元素等同于系数在这个环内的多项式, 等等).

为此, 考虑对于所有 $a \in \mathrm{K}$, 由

$$j(a) = \theta(a, 1)$$

给定的映射

$$j : \mathrm{K} \to \mathrm{F/R}.$$

它是单射的, 因为 $\theta(a', 1) = \theta(a'', 1)$, 可以写成 $a' \cdot 1 = 1 \cdot a''$ 即 $a' = a''$. j 还是环的同态. 事实上我们有

$$\begin{aligned}
j(a') + j(a'') &= \theta(a', 1) + \theta(a'', 1) = \theta(a' \cdot 1 + 1 \cdot a'', 1 \cdot 1) \\
&= \theta(a' + a'', 1) = j(a' + a''), \\
j(a')j(a'') &= \theta(a', 1)\theta(a'', 1) = \theta(a'a'', 1 \cdot 1) = \theta(a'a'', 1) = j(a'a''), \\
j(1) &= \theta(1, 1) = 1.
\end{aligned}$$

于是 j 是从 K 到 F/R 的一个子环的同构, 这个子环由形式为具有 $b = 1$ 的元素 $\theta(a, b)$ 组成. 以后我们在 K 的元素 a 和 F/R 的元素 $\theta(a, 1)$ 之间不做任何区别, 即我们对于任何 $a \in \mathrm{K}$ 写出

$$\theta(a, 1) = a. \tag{10}$$

于是有下列结果: F/R 的所有元素 $\theta(a, b)$ 是 K 的两个元素的商. 更确切地说, 我们有

$$\theta(a, b) = ab^{-1}, \tag{11}$$

或如果愿意写成

$$\theta(a, b) = \theta(a, 1) \cdot \theta(b, 1)^{-1}. \tag{12}$$

注意到由于 $b \neq 0$, $\theta(b, 1)$ 不是零元, 故根据前一小节的结果, 它在 F/R 是可逆的.

为了验证 (12), 只需指出

$$\theta(a, b)\theta(b, 1) = \theta(a, 1),$$

即

$$\theta(ab, b) = \theta(a, 1);$$

而这个关系可以写成 $ab \cdot 1 = ba$, 这显然是成立的.

前面已经构造的域 F/R 称为**交换整环 K 的分式域**. 以后读者可以忘记我们得到它的方式, 对于所有应用只需知道下列性质: K 是它的分式域的子环, 而该分式域的所有元素是 K 的两个元素的商. 尤其是我们再也不使用记号 $\theta(a, b)$ 表示 K 的分式域的元素, 将用记号 a/b 或 $\dfrac{a}{b}$ 或 ab^{-1} 表示它们. 但是读者务必警惕犯这样的本质性错误, 即认为两个分式 a/b 和 c/d 相等, 必须 $a = c$ 并且 $b = d$; 一个分式不是 K 的元素的一个序偶 (a, b), 其中的 $b \neq 0$, 而是这种序偶的一个类.

从有理整数环 $K = \mathbf{Z}$ 出发, 从前面的考虑导出有理数域 $\mathbf{Q}$. 此外人们如果有许多从整数出发定义有理数的 "简单的" 或 "几何的" 方法, 那么仅仅有一个数学上是正确的, 这就是我们所陈述的方法. 假定 $K = \mathbf{Z}$ 丝毫不能简化它 (参见上面的注 1).

5. 系数在一个域内的有理分式

设 K 是一个交换域. 我们已经知道 (§27, 定理 1) 对于所有整数 $n \geqslant 1$, 系数在 K 内的 n 个未定元的多项式环 $K[X_1, \cdots, X_n]$ 是整环. 于是可以应用前面几小节的考虑. 环 $K[X_1, \cdots, X_n]$ 的分式域用记号

$$K(X_1, \cdots, X_n)$$

表示, 并且称为域 K 上的 n **个未定元的有理分式域**. $K(X_1, \cdots, X_n)$ 的元素称为系数在**域 K 内的 n 个未定元的有理分式**.

对于有理分式的实际使用, 求助第二小节的考虑是仍然无用的: 这些考虑只能用来确立域 $K(X_1, \cdots, X_n)$ 的存在性, 而在实际中所关心的是这个域的性质. 换句话说, 读者应当记住下列断言, 其他概不需要:

(FR1) *系数在 K 内的 n 个未定元的有理分式是一个记作 $K(X_1, \cdots, X_n)$ 的域的元素;*

(FR2) *在系数在 K 内的 n 个未定元的有理分式中有系数在 K 内的 n 个未定元的多项式*; 更确切地说, 系数在 K 内的 n 个未定元的多项式环 $K[X_1, \cdots, X_n]$ 是域 $K(X_1, \cdots, X_n)$ 的一个子环;

(FR3) *对于所有的有理分式 $f \in K(X_1, \cdots, X_n)$, 存在两个多项式 $p, q \in K[X_1, \cdots, X_n]$, 这里 $q \neq 0$, 使得有*

$$f = pq^{-1} = p/q.$$

不言而喻, 有多种方式把 f 写成两个多项式的商, 而两个有理分式 p'/q' 和 p''/q'' 是相等的, 必须并且只需 $p'q'' = p''q'$.

例 1 系数在一个域 K 内的一个未定元的有理分式是形如

$$f = \frac{p(X)}{q(X)}$$

的表达式, 其中 p 和 q 是系数在 K 内的一个未定元的多项式, $q \neq 0$. 例如

$$f = \frac{X^3 + X^2 - 1}{X^2 - 1}.$$

注意 $q \neq 0$ 不排除对于某些 $x \in K$ 有 $q(x) = 0$ 的可能性. 它简单地表示 q 不是环 $K[X]$ 的元素 0, 即它的系数不全是零.



例 2　表达式

$$f = \frac{X + Y}{X - Y}$$

是系数在 $\mathbf{Q}$ 内的两个未定元的有理分式.

6. 有理分式的值

设 L 是一个交换域, K 是 L 的一个子域, 而 f 是一个系数在 K 内的 n 个未定元的有理分式. 设

$$u = (u_1, \cdots, u_n)$$

是 L^n 的一个元素. 说 *f 在 u 有定义*, 或 *u 是可代入 f 的*, 如果存在多项式 p 和 q, 使得

$$f = p/q, \quad q(u_1, \cdots, u_n) \neq 0;$$

这并非说任意使得 $f = p/q$ 的多项式 p 和 q $(q \neq 0)$ 都必须有 $q(u_1, \cdots, u_n) \neq 0$.

例 3　$u = 0$ 是可以代入有理分式

$$f = \frac{X^2}{X^2 + X}$$

的, 因为它还可以写成

$$\frac{X}{X+1},$$

而在这种形式下, 我们看到对于 $u = 0$ 分母不是零.

假定 f 在 $(u_1, \cdots, u_n)$ 有定义, 则可以用下列方式定义 *f 在 u 的值*:

$$f = p/q, \quad 其中\ q(u) \neq 0,$$

L 的元素 $p(u)/q(u)$ 仅依赖 f, 而不依赖 p 和 q 的选取. 事实上, 如果还有 f 的另一个表示

$$f = p'/q', \quad 其中\ q'(u) \neq 0,$$

则有

$$pq' = p'q,$$

故 (§28, 公式 (6))

$$p(u)q'(u) = p'(u)q(u),$$

因此有

$$p(u)/q(u) = p'(u)/q'(u),$$

这就证实了我们的断言. 明确了这一点, L 的元素 $p(u)/q(u)$ 就称为 f 在 u 的值, 用记号

$$f(u) \quad 或 \quad f(u_1, \cdots, u_n)$$

表示这个值, 于是有

$$f(u)=p(u)/q(u),\quad \text{如果 } f=p/q \text{ 并且 } q(u)\neq 0.$$

显然如果 f 是多项式, 那么 f 对于任意 $u\in \mathrm{L}^n$ 都有定义, 并且刚才定义的值 $f(u)$ 跟 §28 第 1 小节所定义的值一致. 为了看清这一点, 只需令 $f=f/1$.

我们指出在一个给定的点 $u\in \mathrm{L}^n$ 有定义的有理分式的集合是 $\mathrm{K}(\mathrm{X}_1,\cdots,\mathrm{X}_n)$ 的一个子环. 事实上, 假定 f' 和 f'' 在 u 有定义, 那么可以写出

$$f'=p'/q',\quad \text{其中 } q'(u)\neq 0,$$
$$f''=p''/q'',\quad \text{其中 } q''(u)\neq 0,$$

由此得到

$$f'+f''=\frac{p'q''+p''q'}{q'q''},\quad f'f''=\frac{p'p''}{q'q''},$$

由于有 $q'(u)q''(u)\neq 0$, 我们发现 $f'+f''$ 和 $f'f''$ 在 u 有定义. 由于多项式 (其特殊情形是 1) 在 u 有定义, 因此在 u 有定义的有理分式必然组成 $\mathrm{K}(\mathrm{X}_1,\cdots,\mathrm{X}_n)$ 的一个子环.

假定 f' 和 f'' 在 $u\in \mathrm{L}^n$ 有定义, 那么分式 $f'+f''$ 和 $f'f''$ 在 u 的值是 $f'(u)+f''(u)$ 和 $f'(u)f''(u)$. 事实上, 保留前面的记号, 设 $g=f'+f''$, 那么就有

$$g=p/q,\quad \text{其中 } p=p'q''+p''q',\ q=q'q'',$$

于是

$$g(u)=p(u)/q(u)=\frac{p'(u)q''(u)+p''(u)q'(u)}{q'(u)q''(u)}=\frac{p'(u)}{q'(u)}+\frac{p''(u)}{q''(u)},$$

这就验证了所宣布的第一个结果, 第二个的验证类似.

假定 f 在 u 有定义, 那么 f^{-1} 在 u 有定义必须并且只需 $f(u)\neq 0$, 并且有

$$f^{-1}(u)=f(u)^{-1}.$$

事实上, 假定 f 和 f^{-1} 在 u 有定义, 那么有 $1=f\cdot f^{-1}$, 由此在 u 取值得到

$$1=f(u)\cdot f^{-1}(u),$$

这就验证了 $f(u)\neq 0$, 并且 $f^{-1}(u)=f(u)^{-1}$. 反之, 假定 f 在 u 有定义, 并且 $f(u)\neq 0$, 那么有 $f=p/q$, 其中 $q(u)\neq 0$, 并且由于 $f(u)\neq 0$, 还有 $p(u)\neq 0$. 写出 $f^{-1}=q/p$, 则得到 f^{-1} 在 u 也有定义.

例 4 取 $n=2$ 和有理分式

$$f=\frac{\mathrm{X}+\mathrm{Y}}{\mathrm{X}-\mathrm{Y}},$$

它在所有满足 $u \neq v$ 的 $(u,v) \in \mathrm{L}^2$ (即对角线以外) 有定义, 不存在 f 有定义的任何其他的点. 事实上, 设 p 和 q 是系数在 K 内的多项式, 使得 $f = p/q$, 于是有

$$(\mathrm{X}+\mathrm{Y})q(\mathrm{X},\mathrm{Y}) = (\mathrm{X}-\mathrm{Y})p(\mathrm{X},\mathrm{Y});$$

令 $p = \sum p_n, q = \sum q_n$, 这里 p_n 和 q_n 是总次数为 n 的齐次多项式, 直接从前边的关系得到

$$(\mathrm{X}+\mathrm{Y})q_n(\mathrm{X},\mathrm{Y}) = (\mathrm{X}-\mathrm{Y})p_n(\mathrm{X},\mathrm{Y}).$$

由此推出 (读者作为习题通过展示 p_n 和 q_n 的系数并且建立这些系数之间的关系将证明它), 对于所有的 n 存在一个多项式 $r_n(\mathrm{X},\mathrm{Y})$, 使得

$$q_n(\mathrm{X},\mathrm{Y}) = (\mathrm{X}-\mathrm{Y})r_n(\mathrm{X},\mathrm{Y}).$$

因此对于 $u = v$ 显然有 $q(u,v) = 0$, 故 f 在 L^2 的对角线的任何点均未定义.

例 5 K 是任何一个交换域, 取

$$\mathrm{L} = \mathrm{K}(\mathrm{X}_1, \cdots, \mathrm{X}_n),$$

取

$$u = (\mathrm{X}_1, \cdots, \mathrm{X}_n) \in \mathrm{L}^n.$$

那么所有有理分式 $f \in \mathrm{K}(\mathrm{X}_1, \cdots, \mathrm{X}_n)$ 在 u 有定义, 这是因为当写出 $f = p/q$ 时, 其中 $q \neq 0$, 由于像在 §28 注 1 中曾经看到的那样

$$q(\mathrm{X}_1, \cdots, \mathrm{X}_n) = q,$$

因此可以定义 $f(\mathrm{X}_1, \cdots, \mathrm{X}_n)$. L 的这个元素由

$$f(\mathrm{X}_1, \cdots, \mathrm{X}_n) = p(\mathrm{X}_1, \cdots, \mathrm{X}_n)/q(\mathrm{X}_1, \cdots, \mathrm{X}_n) = p/q = f$$

给定. 这就解释了在实际中为什么经常用记号 $f(\mathrm{X}_1, \cdots, \mathrm{X}_n)$ 表示一个有理分式. 这里所进行的讨论使得能够证明 $f(\mathrm{X}_1, \cdots, \mathrm{X}_n) = f$, 而在经典的教材里, 这个关系仅作为书写的一种简单约定.

这个例子 (以及许多其他的例子) 解释了为什么定义一个系数在 K 内的有理分式在其坐标不在 K 内而在 K 的任意一个扩张的一个点取值是必要的. 在最初等的实践中, 显然有必要对于实数系数的有理分式在坐标为复数的点赋值.



注 2 如果 L 是 K 的一个扩张, 而有理分式 $f \in \mathrm{K}(\mathrm{X}_1, \cdots, \mathrm{X}_n)$ 在点 $u \in \mathrm{L}^n$ 没有定义时, 两种情形是可能的: 一种是 f 的逆元 $1/f$ 在 u 有定义 (这时说 u 是 f 的**极点**), 一种是 $1/f$ 在 u 没有定义 (这时说 u **是** f 的**不定点**). 当且仅当可以写出

$$f = p/q, \quad \text{其中 } p(u) \neq 0,\ q(u) = 0$$

第一种情形发生. 如果是这样, 那么 $f^{-1}=q/p$ 显然在 u 有定义, 并且在 u 取值为 0, 以致 f 在 u 不能有定义; 反之, 如果 u 是 f 的一个极点, 因为 f^{-1} 在 u 有定义, 可以写出

$$f^{-1}=q/p,\quad 其中\ p(u)\neq 0;$$

如果有 $f^{-1}(u)\neq 0$, f 像上面看到的那样将在 u 也有定义, 由假设这种情形不会出现, 于是有 $f^{-1}(u)=0$, 即 $q(u)=0$, 因此正如所断言的那样有

$$f=p/q,\quad 其中\ p(u)\neq 0,\ q(u)=0.$$

上述考虑还允许刻画 f 的不定点: u 是 f 的不定点, 如果对于任意满足条件

$$f=p/q,\quad 其中\ q\neq 0$$

的多项式 p 和 q 有

$$p(u)=q(u)=0.$$

例如取 $\mathrm{K}=\mathrm{L}=\mathbf{C}$ 和 f 是两个未定元的有理分式

$$f=\frac{\mathrm{X}+\mathrm{Y}}{\mathrm{X}-\mathrm{Y}};$$

它显然在所有使得 $u\neq v$ 的点 $(u,v)\in\mathbf{C}^2$ 有定义. 异于 $(0,0)$ 的对角线 $u=v$ 的点显然是 f 的极点; 最后, 点 $(0,0)$ 是 f 的不定点. 为了建立这后一个结果, 必须证明, 如果两个多项式 $p,q\in\mathbf{C}[x,y]$ 满足

$$\frac{p}{q}=\frac{\mathrm{X}+\mathrm{Y}}{\mathrm{X}-\mathrm{Y}},$$

则必然有

$$p(0,0)=q(0,0)=0.$$

令

$$p=a+a'\mathrm{X}+a''\mathrm{Y}+\cdots,$$
$$q=b+b'\mathrm{X}+b''\mathrm{Y}+\cdots,$$

没有写出的项的次数至少是 2. 由假设

$$(\mathrm{X}-\mathrm{Y})p=(\mathrm{X}+\mathrm{Y})q,$$

即

$$(\mathrm{X}-\mathrm{Y})(a+a'\mathrm{X}+a''\mathrm{Y}+\cdots)=(\mathrm{X}+\mathrm{Y})(b+b'\mathrm{X}+b''\mathrm{Y}+\cdots),$$

因此有

$$a(X-Y)=b(X+Y), \quad 即 \quad a=b=-b,$$

由此显然得到 $a=b=0$. 由于 $p(0,0)=a$ 和 $q(0,0)=b$, 我们的断言得以证明.

有理分式以及类似对象 (“多变量代数函数”) 的研究是代数几何的主要目的之一.

§29 习题

¶1. 考虑两个有理分式 $f,g\in K(X_1,\cdots,X_n)$, 其中 K 是一个*无限域*. 假定在 K^n 内存在非空的 Zariski 开集 A (§§27, 28, 习题 1), 使得 f 和 g 在所有的点 $x\in A$ 有定义并且取同样的值. 证明那么就有 $f=g$. (这个结果, 特别在 $K=\mathbf{R}$ 或 $\mathbf{C}$ 的经典情形, 解释了为什么可以把系数在 K 内的有理分式等同于定义在 K^n 的一个子集上的函数.)

2. 一个变量的一个有理函数不具有任何不定点 (如果 $f=p/q$, 其中的 p 和 q 同时在 a 取零值, 分解出在 p 和 q 内的 $X-a$ 的最高次幂).

3. 设 K 是一个交换域, a 是 K 的一个元素, 而 A 是在 a 有定义的有理分式 $f\in K(X)$ 的集合. 证明 A 是域 K[X] 的一个赋值环 (§8, 习题 6), 并且 A 的非可逆元的理想由使得 $f(a)=0$ 的 $f\in A$ 组成.

证明可以写成形式 $f=p/q$ 的元素 $f\in K[X]$, 其中的 p 和 q 是满足条件

$$d^\circ(p)\leqslant d^\circ(q)$$

的多项式, 同样组成 K[X] 的赋值环.

4. 设 L 是一个交换域, K 是 L 的一个子域, 而 $x_1,\cdots,x_n$ 是 L 的元素. 用 $K(x_1,\cdots,x_n)$ 表示含有 K 和各个 x_i 的 L 的最小子域 (称为**由 K 和 x_i 生成的 L 的子域**; 包含 K 作为子域并且由 K 和有限个元素生成的一个域称为 K 上的**代数函数域**). 证明这是 L 所有可以写成形式

$$f(x_1,\cdots,x_n)$$

的元素的集合, 这里 $f\in K(X_1,\cdots,X_n)$ 在 $(x_1,\cdots,x_n)$ 有定义. 证明 $K(x_1,\cdots,x_n)$ 同构于 $K(X_1,\cdots,X_n)$, 如果 x_i 在 K 上是代数无关的.

¶5. 设 K 是一个交换域. 给定一个不在 K 内的有理分式 $f\in K(X)$, 证明域 K(X) 的元素 X 在由 K 和 f 生成的子域 $K(f)$ 上是代数的. 由此推出所有的 $g\in K(X)$ 亦如此.

*证明给定两个多项式 $p,q\in K[X]$, 则存在 p 和 q 之间的一个非平凡的系数在 K 内的***代数关系**.

¶¶6. 设 L 是一个交换域, 而 K 是 L 的一个子域.

a) 设 $x_1,\cdots,x_r,y,z$ 是 L 的元素, 假定 z 在子域 $K(x_1,\cdots,x_r,y)$ 上是代数的, 但在 $K(x_1,\cdots,x_r)$ 上不是代数的. 证明 y 在

$$K(x_1,\cdots,x_r,z)$$

上是代数的.

b) 假定 L 是在 K 上**有限超越次的**, 即存在一个整数 n, 使得 L 的任意 $n+1$ 个元素满足一个系数在 K 内的非平凡的代数关系. 证明那么就可以找到 L 的有限个在 K 上代数无关的元素

$x_1, \cdots, x_r$, 使得 L 的所有元素在子域 $K(x_1, \cdots, x_r)$ 上是代数的 (这时就说 x_i 组成在 K 上 L 的**超越基**).

c) 设 $x_1, \cdots, x_r$ 和 $y_1, \cdots, y_s$ 是 K 上的 L 的两个超越基. 证明存在一个指标 j, 使得 y_j 在 $K(x_1, \cdots, x_{r-1})$ 上不是代数的 (相反, L 的所有元素, 特别是 x_r, 在 $K(x_1, \cdots, x_{r-1})$ 上是代数的). 借助问题 a), 由此推出 $x_1, \cdots, x_{r-1}, y_j$ 组成在 K 上 L 的一个超越基.

d) 由上推出在 K 上 L 的任意两个超越基有同样的元素数 (称为在 K 上 L 的**超越次数**). 证明这个数是使得满足在 K 上代数无关的 L 的 n 个元素的 n 中的最大者.

e) 证明, 如果 $f_1, \cdots, f_{n+1}$ 是 $n+1$ 个系数在交换域 K 内的 n 个未定元的有理分式, 则存在 $f_1, \cdots, f_{n+1}$ 之间的系数在 K 内的非平凡代数关系.

f) 假定交换域 K 是无限的. 设 A 是在 K^p 内的一个非空 Zariski 开集 (§§27, 28, 习题 1). 说从 A 到 K^q 内的一个映射 f 是**有理的**, 如果存在有理分式

$$f_1, \cdots, f_q \in K(X_1, \cdots, X_p),$$

使得

$$f(x) = (f_1(x), \cdots, f_q(x)) \quad \text{对于所有 } x \in A.$$

(这个概念推广了 §§27, 28 的习题 17 的多项式映射的概念.) 有了这个定义, 证明, 如果从 A 到 K^q 内的一个有理的映射 f 是满射的, 则 $p \geqslant q$. (这个结果表明, 如果限制到由多项式或有理分式定义的映射, 则像 Peano 曲线那样的 "病态的" 现象 —— 存在从直线到平面上的一个连续映射 —— 就不会发生.)

[这个习题里所陈述的超越次数的概念是代数流形维数定义的基础.

设 V 是 $\mathbf{C}^n$ 的一个**代数流形**, 即由有限个方程

$$f_1(x) = \cdots = f_r(x) = 0$$

的方程组所定义的 $\mathbf{C}^n$ 的一个子集, 其中的 $f_1, \cdots, f_r$ 是系数在 $\mathbf{C}$ 内的 n 个变量的多项式. 从 V 到 $\mathbf{C}$ 内的所有映射, 如果它是 $\mathbf{C}^n$ 上的多项式函数在 V 上的限制, 则称为 **V 上的多项式函数**. V 上的这些多项式函数显然组成包含 $\mathbf{C}$ (常值函数) 的一个环 A, 并且是由适当选取的 n 个元素 (例如 $\mathbf{C}^n$ 的坐标函数在 V 上的限制) 生成的. 称 V 是**不可约的**, 如果环 A 是整环; 或同样的, 正如可以证明的结果所说的, 这就是要求 V 不能够写成与 V 不同的其他的代数流形的并集. 如果 V 是不可约的, 则可以组成 A 的分式的域 L. 记 $\mathbf{C}^n$ 的坐标函数在 V 上的限制为 $f_1, \cdots, f_n$, 则显然有

$$L = \mathbf{C}(f_1, \cdots, f_n).$$

称 L 是**流形 V 的有理函数域**. 说了这些之后, 那么 L 就是在 $\mathbf{C}$ 上**有限超越次**的, 称 L 在 $\mathbf{C}$ 上的超越次数为 V 的**维数**. 一条 "曲线" 是一维的, 一张 "曲面" 是二维的, 等等. 已经证明, 一个不可约代数流形 V 的维数 p 是最大的整数, 它使得可以构造一个非空的并且两两不同的不可约代数流形的递增链

$$V_0 \subset V_1 \subset \cdots \subset V_p = V.$$

一张曲面包含一条曲线, 而一条曲线含有一个点, 这或许就解释了为什么一张曲面是二维的.

在 $\mathbf{C}^n$ 内的代数曲面用在分析中, 尤其是在常系数线性偏微分方程的研究中. 例如考虑方程

$$\sum_{i_1\cdots i_n\geqslant 0} a_{i_1\cdots i_n}\frac{\partial^{i_1+\cdots+i_n}f}{\partial x_1^{i_1}\cdots\partial x_n^{i_n}}=0,$$

其中的 f 是 n 个变量的未知函数, 系数 $a_{i_1\cdots i_n}$ 是几乎全部为零的常数. 如果要找形式为

$$f(x_1,\cdots,x_n)=e^{u_1x_1+\cdots+u_nx_n}$$

的解, 其中 $u_1,\cdots,u_n$ 是复常数, 这显然就归结为解方程

$$\sum a_{i_1\cdots i_n}u_1^{i_1}\cdots u_n^{i_n}=0,$$

即研究由这个方程定义的 $\mathbf{C}^n$ 中的超曲面.]

7. 求下列有理分式的极点和不定点:

$$\frac{\mathrm{X}}{\mathrm{Y}};\quad \frac{\mathrm{X-Y}}{\mathrm{XY}};\quad \frac{(\mathrm{X}^2-1)(\mathrm{Y}^2-1)}{\mathrm{X}^2+\mathrm{Y}^2-1};\quad \frac{\mathrm{X+Y+Z}}{\mathrm{X-Y}};\quad \frac{\mathrm{X-Z}}{\mathrm{Y-Z}}$$

(取 $\mathbf{C}$ 为基域).

¶8. 设 K 是一个交换域. 考虑 (§§27, 28, 习题 11) 系数在 K 内的一个未定元的形式幂级数环 K[[X]]. 由于这是一个整环, 可以组成它的分式域, 记为 K((X)). 证明这个分式域的所有元素以唯一的方式写成如下形式: X 的一个幂 (可能有负指数) 乘以一个形式级数, 其常项不是零, 即是一个系数在 K 内的 "级数"

$$\sum_{n=-\infty}^{\infty} a_n\mathrm{X}^n, \tag{$*$}$$

其中的使得 $a_n\neq 0$ 的负整数 n 的个数有限.

我们注意到, 像 K[X] 是 K[[X]] 的一个子环一样, 域 K((X)) 包含一个同构于 K(X) 的子域, 以致系数在 K 内的一个未定元的有理分式可以用形如 (∗) 的一个级数表示. 求表示下列有理分式的形式幂级数 (∗):

$$\frac{1}{\mathrm{X}-\mathrm{X}^2};\quad \frac{\mathrm{X}^2+\mathrm{X}+1}{\mathrm{X}^4-\mathrm{X}^2}.$$

证明表示有理分式 f 的形式幂级数 (∗) 不包括 X 的任何负指数幂, 如果 f 在 $x=0$ 有定义, 反之亦真(∗).

证明 K[[X]] 是域 K((X)) 的一个赋值环 (§8, 习题 6).

¶9. 设 A 是一个交换环, 而 S 是 A 的一个子集. 假定 S 含有 1 但是不含有 0, 并且对于任意 $x\in\mathrm{S}, y\in\mathrm{S}$ 有 $xy\in\mathrm{S}$ (如果 A 是一个整环, 举例说, 可以取 A 的非零元素的集合; 在一般情形, 一个重要的例子是取不属于 A 的一个给定的素理想的元素的集合作为 S).

(∗) 给定了两个多项式 $f,g\in\mathrm{K[X]}$, 把有理分式 f/g 写成形如 (∗) 的形式幂级数, 这种运算在老的代数书里称为按照 X 的升幂 f 除以 g; 这个运算在于, 对于每个整数 $r\geqslant 0$ (如果 g 的常值项非零, 很明显总可以归结到这种情形) 求次数 $\leqslant r$ 的多项式 q, 使得 $f(\mathrm{X})-q(\mathrm{X})g(\mathrm{X})$ 是 X^{r+1} 的倍式. 从表示 f/g 的形式幂级数 (∗) 去掉次数 $>r$ 的项即得 q.

在实际中, 当要展开两个多项式或幂级数的商为幂级数时, 或当求有限展开时, 就应当进行这样的运算.

a) 设 F 是这样的序偶 (x,s) 的集合, 其中 $x \in \mathrm{A}, s \in \mathrm{S}$. 给定 F 的两个元素 $y'=(x',s')$ 和 $y''=(x'',s'')$, 用 $\mathrm{R}\{y',y''\}$ 表示关系

$$\text{存在 } s \in \mathrm{S}, \quad \text{使得 } s(x's''-x''s')=0.$$

证明 R 是 F 上的一个等价关系. 当 A 是一个整环, 并且 S 是 A 的非零元素的集合时, 将会发生什么?

b) 设 $\mathrm{A_S}$ 是商集 F/R, 而 θ 是从 F 到 F/R 上的典范映射. 证明在 $\mathrm{A_S}$ 上存在唯一的一个环结构, 使得有公式

$$\theta(x,s)+\theta(y,t)=\theta(xt+ys,st),$$
$$\theta(x,s)\cdot\theta(y,t)=\theta(xy,st).$$

c) 证明由

$$j(x)=j(x,1)$$

定义的从 A 到 $\mathrm{A_S}$ 内的映射是环的一个同态, 对于所有的 $s \in \mathrm{S}$, $j(s)$ 是可逆的, 并且 $\mathrm{A_S}$ 的所有元素是一个元素 $j(x)(x \in \mathrm{A})$ 除以一个元素 $j(s)(s \in \mathrm{S})$ 的商. 同态 j 的核是什么? 在什么情形下它是单射?

d) 设 f 是从 A 到一个交换环 K 内的一个同态. $f(s)$ 对于任意 $s \in \mathrm{S}$ 都是可逆的, 必须并且只需 f 是上一个问题中的 j 和从环 $\mathrm{A_S}$ 到环 K 内的一个同态的复合.

e) 设 A 是一个交换整环, K 是它的分式域, 而 $\mathfrak{p}$ 是 A 的一个素理想 (即 $\mathfrak{p} \neq \mathrm{A}$, 并且关系 $xy \in \mathfrak{p}$ 蕴含 $x \in \mathfrak{p}$ 或 $y \in \mathfrak{p}$). 取 $\mathfrak{p}$ 在 A 内的补集为 S. 证明环 $\mathrm{A_S}$ 同构于由可以写成形式 $x/y(x,y \in \mathrm{A}, y \notin \mathfrak{p})$ 的分式组成的 K 的子环 $\mathrm{A}_\mathfrak{p}$.

f) 保留 e) 的假设, 令环 $\mathrm{A}_\mathfrak{p}$ 的每一个异于 $\mathrm{A}_\mathfrak{p}$ 的理想 $\mathfrak{a}$ 对应环 A 的理想 $\mathrm{A} \cap \mathfrak{a}$. 证明以这种方式就得到了一个从 $\mathrm{A}_\mathfrak{p}$ 的异于 $\mathrm{A}_\mathfrak{p}$ 的理想的集合到包含在 $\mathfrak{p}$ 内的 A 的理想的集合上的一个双射. 逆映射是什么?

g) 设 L 是一个交换域, 而 $a_1,\cdots,a_n$ 是 L 的元素. 取

$$\mathrm{A}=\mathrm{L}[\mathrm{X}_1,\cdots,\mathrm{X}_n],$$

而取使得 $f(a_1,\cdots,a_n)=0$ 的多项式的集合为 $\mathfrak{p}$. 证明 $\mathrm{A}_\mathfrak{p}$ 是在 L^n 的点 $(a_1,\cdots,a_n)$ 有定义, 系数在 L 内有理分式 $f(\mathrm{X}_1,\cdots,\mathrm{X}_n)=0$ 的集合.

h) 推广问题 f) 的结果到任意分式环 $\mathrm{A_S}$ 的情形.

10. 设 A 是一个交换整环, 而 K 是它的分式域. 证明多项式环 A[X] 的分式域典范地同构于有理分式域 K(X).

¶11. 设 A 是一个交换整环, M 是一个 A-模, 而 K 是 A 的分式域. 我们打算证明, 如果 M 是*无扭的* (§10, 习题 11, 这里用不到这个习题的结果), 则可以把 M 嵌入到 K 上的一个向量空间 (平凡的例子: A^n 可以嵌入到 K^n).

关于 M 不做任何假设, 直到有新的假设提出.

a) 设 F 是这样的序偶 (m,s) 的集合, 其中 $m \in \mathrm{M}, s \in \mathrm{A}$, 并且 $s \neq 0$. 给定 F 的两个元素 $x'=(m',s')$ 和 $x''=(m'',s'')$, 用 $\mathrm{R}\{x',x''\}$ 表示关系

$$\text{存在 } s \in \mathrm{A}, \quad \text{使得 } s(s'm''-s''m')=0, \quad \text{并且 } s \neq 0.$$

证明 R 是集合 F 上的一个等价关系.

b) 设 V 是商集 F/R, 而 θ 是从 F 到 V 上的典范映射. 证明可以定义 V 的两个元素的和, 使得对于任意 (m', s') 和 $(m'', s'') \in \mathrm{F}$ 有

$$\theta(m', s') + \theta(m'', s'') = \theta(s''m' + s'm'', s's''),$$

并且配备了这个运算的 V 是一个交换群.

c) 证明存在一个从 $\mathrm{K} \times \mathrm{V}$ 到 V 内的映射 $(\lambda, x) \to \lambda x$, 它满足下列条件: 如果 $\lambda = u/s$ 并且 $x = \theta(m, t)$ (其中的 $u, s, t \in \mathrm{A}, m \in \mathrm{M}$, 并且 s, t 非零), 则有

$$\lambda x = \theta(um, st).$$

证明配备了这个运算的交换群 V 是 K 上的一个向量空间.

d) 定义从 M 到 V 内的一个 "典范" 映射

$$j(m) = \theta(m, 1).$$

证明 j 是 A-模的一个同态 (注意: 由于 V 是 K 上的一个向量空间, 更加可以把 V 看作一个 A-模), 其核是 M 的扭子模 (即这样的 m 的集合, 至少对于一个非零的 $s \in \mathrm{A}$ 有 $sm = 0$). 由此推出, 如果 M 是无扭的, 则 j 是从 M 到 V 的一个子模上的同构. 例子 (当取 $\mathrm{A} = \mathbf{Z}$ 时): 所有无扭交换群嵌入到一个有理向量空间.

e) 假定 M 是有限生成的和无扭的. 证明 V 在 K 上是有限维的. 设 $n = \dim(\mathrm{V})$ (称 n 是 M 的**秩**), 证明存在 V 的两个基 $(a_i)_{1\leqslant i\leqslant n}$ 和 $(b_i)_{1\leqslant i\leqslant n}$, 使得如果用 P 和 Q 分别表示由 a_i 和 b_i 生成的 V 的子 A-模 (同构于 A^n), 那么有 $\mathrm{P} \subset \mathrm{M} \subset \mathrm{Q}$.

f) 由此和 §18 的定理 3 推出下列结果: 如果 A 是一个主理想整环, 则所有无扭的和有限生成的 A-模 M 同构于 A^n, 这里 n 是 M 的秩. 当 $\mathrm{A} = \mathbf{Z}$ 时叙述这个结果.

g) 设 M 是主理想整环 A 上的一个有限生成的模. 设 T 是 M 的扭子模. 证明 M/T 是有限生成和自由的. 由此推出 (借助 §17 的习题 8) M 同构于 T 和一个有限生成的自由的 A-模的直积. (这个结果把有限生成模的研究归结为有限生成扭模的研究. 在 §31, 习题 8, 9, 10 将会这样做.)

h) 设 M 是主理想整环 A 上的一个有限生成的自由的模, 而 M' 是 M 一个子模. 证明下列性质是等价的: i) M' 是 M 的直和项; ii) 商模 M/M' 是无扭的; iii) 对于任意的 $a \in \mathrm{A}$ 和 $x \in \mathrm{M}$, 关系 $ax \in \mathrm{M}'$ 蕴含 $a = 0$ 或 $x \in \mathrm{M}'$. 由此重新获得 §18 习题 2 的结果.

i) 设 M' 是由整系数的齐次线性方程组定义的 $\mathbf{Z}^n$ 的一个子群. 证明 M' 的所有的基是 $\mathbf{Z}^n$ 的一个基的子集.

12. 在一本面向中学生的代数课本里, 人们读到这样的句子: "不让变量取使分子或分母变为零的值, 配备了加法和乘法运算的有理分式集合就具有域结构. " 对于这种陈述你认为如何?

§30 导子和 Taylor 公式

在实变量函数的理论中, 一个函数 $f(t)$ 的导数 $f'(t)$ 是借助 "取极限" 即完全不用代数过程定义的. 然而, 当 f 是一个多项式函数时, 设

$$f(t) = \sum a_r t^r,$$

则 f 的导数仍然是多项式函数, 其导数是

$$f'(t) = \sum ra_r t^{r-1}.$$

显然如果限于研究多项式函数, 可以借助上述公式定义一个函数的导数, 在这样的公式里, 不出现任何取极限的运算.

这些评注暗示了推广导数概念到系数在任意的交换环内的多项式的可能性, 并且有可能通过纯代数的手段证明其性质, 这些性质在经典的情形通过“分析”推理得到, 这种推理可以用于比多项式函数广泛得多的函数. 本节要做的就是用代数方法引进多项式函数导数的定义并证明其性质. 这样得到的结果不单纯是古典理论的优美却缺乏实际用处的推广, 后面为了区别代数方程的单根和重根就要实际用到它们, 并且今日被用到其他重要问题 (例如, 为了区别代数流形的“单点”和“重点”).

1. 环的导子

设 K 是一个环, 称所有的映射

$$\mathrm{D} : \mathrm{K} \to \mathrm{K}$$

为**环 K 的导子**, 如果它对于任意 $x, y \in \mathrm{K}$ 满足关系

$$\mathrm{D}(x+y) = \mathrm{D}(x) + \mathrm{D}(y), \tag{1}$$

$$\mathrm{D}(xy) = \mathrm{D}(x)y + x\mathrm{D}(y). \tag{2}$$

这个定义显然是受到了导数的经典计算法则的启示.

例 1 取 K 为 $\mathbf{R}$ 上的多项式环. 如果

$$x(t) = \sum a_r t^r$$

是一个这样的函数, 定义 $\mathrm{D}(x)$ 为函数

$$x(t) = \sum ra_r t^{r-1},$$

我们所熟悉的和与乘积的导数的古典法则就表明这样定义的映射是环 K 的导子.

关系 (1) 和 (2) 表明有

$$\mathrm{D}(1) = 0. \tag{3}$$

因为在 (2) 中令 $y=1$ 则得对于所有 x 有 $x \cdot \mathrm{D}(1) = 0$, 特别取 $x=1$ 则得 $\mathrm{D}(1)=0$. 如果 K 是交换的, 则对于所有 $x \in \mathrm{K}$ 和所有整数 $n \geqslant 0$, 更一般的, 有

$$\mathrm{D}(x^n) = nx^{n-1}\mathrm{D}(x). \tag{4}$$

对于 $n=0$ 这个结果事实上归结为 (3), 如果它对于 $n-1$ 成立, 那么 (2) 表明当 K 是交换环时有

$$\begin{aligned}D(x^n) &= D(x^{n-1}\cdot x) = D(x^{n-1})x + x^{n-1}D(x)\\ &= (n-1)x^{n-2}D(x)x + x^{n-1}D(x) = nx^{n-1}D(x).\end{aligned}$$

2. 多项式环的导子

我们要证明下列结果:

定理 1　*设 D 是交换环 K 上的一个导子. 给定一个多项式 $u\in K[X]$, 在环 K[X] 内存在唯一的一个导子在 K 上跟 D 一致, 并且映射多项式 X 到给定的多项式 u.*

假定发现了这样的导子 D′, 给定一个多项式

$$f=\sum a_r X^r \quad (a_r\in K),$$

法则 (1), (2), (4) 给出

$$D'(f)=\sum D'(a_rX^r)=\sum[D'(a_r)X^r + ra_rX^{n-1}D'(X)].$$

考虑到加在 D′ 上的条件即得

$$D'(f)=\sum D(a_r)X^r + u\sum ra_rX^{n-1}. \tag{5}$$

为了表达得到的结果, 引进两个记号是方便的, 一个是把 D 应用到 f 的系数上得到多项式

$$f^{D}(X)=\sum D(a_r)X^r, \tag{6}$$

另一个是 f 的导多项式

$$f'(X)=\sum ra_rX^{r-1}. \tag{7}$$

这样关系 (5) 就可以写成形式

$$D'(f)=f^{D}+u\cdot f', \tag{8}$$

并且验证了 D′ 的唯一性.

还要指出由公式 (8) 定义的映射

$$D': K[X]\to K[X]$$

是满足所宣布的条件的一个导子. 显然的公式

$$(f+g)^{D}=f^{D}+g^{D},\quad (f+g)'=f'+g'$$

已经表明 D' 满足 (1). 为了验证 (2), 首先注意如果 D_1 和 D_2 是环 L 的导子, 那么对于任意 $u_1, u_2 \in L$, 从 L 到 L 内的映射

$$x \to u_1 \cdot D_1(x) + u_2 \cdot D_2(x)$$

还是 L 的导子. 为了指出 D' 满足 (2), 根据 (8) 只需证明在 $K[X]$ 内两个映射

$$f \to f^{D}, \quad f \to f'$$

是导子, 即有公式

$$(fg)^{D} = f^{D} \cdot g + f \cdot g^{D}, \quad (fg)' = f' \cdot g + f \cdot g'.$$

先假定 f 和 g 缩减为单项式

$$f = aX^p, \quad g = bX^q,$$

则有

$$(fg)^{D} = (abX^{p+q})^{D} = D(ab)X^{p+q} = D(a)X^p \cdot bX^q + aX^p \cdot D(b)X^q,$$
$$(fg)' = (abX^{p+q})' = (p+q)abX^{p+q-1} = paX^{p-1} \cdot bX^q + aX^p \cdot qbX^{q-1},$$

这就显然在特殊情形证明了要得到的公式. 在一般情形, 分解 f 和 g 成单项式的和, 这就把问题引导至刚证明的特殊情形.

我们已经证明了映射 (8) 必然是一个导子, 还需一方面验证它在 K 上与 D 一致, 即

$$f = a \in K \quad 蕴含 \quad D'(f) = D(a).$$

根据 (5) 这是显然的. 另一方面验证

$$D'(X) = u,$$

如果写出 $X = 1 \cdot X$, 并且注意到 $D(1) = 0$, 这同样是显然的. 至此定理 1 证明完毕.

3. 偏导子

定理 1 可以推广到多个未定元的多项式如下:

定理 2 设 K 是一个交换环, D 是 K 的一个导子, 而 $u_1, \cdots, u_n$ 是系数在 K 内的 n 个未定元的多项式. 则存在环 $K[X_1, \cdots, X_n]$ 唯一的一个导子 D', 它在 K 上简化为 D, 并且满足

$$D'(X_i) = u_i \quad 对于 \ 1 \leqslant i \leqslant n.$$

证明显然类似于定理 1, 我们仅需指出对于多项式

$$f = \sum a_{r_1\cdots r_n}X_1^{r_1}\cdots X_n^{r_n}$$

如何计算 $D'(f)$. 由于 D' 是一个导子, 我们有

$$\begin{aligned}D'(f) &= \sum D'(a_{r_1\cdots r_n}X_1^{r_1}\cdots X_n^{r_n})\\ &= \sum D'(a_{r_1\cdots r_n})X_1^{r_1}\cdots X_n^{r_n} + \sum a_{r_1\cdots r_n}D'(X_1^{r_1})X_2^{r_2}\cdots X_n^{r_n}\\ &\quad + \cdots + \sum a_{r_1\cdots r_n}X_1^{r_1}\cdots X_{n-1}^{r_{n-1}}D'(X_n^{r_n}).\end{aligned}$$

由于 $D'(a) = D(a), D'(X_i) = u_i$, 故得

$$\begin{aligned}D'(f) = &\sum D(a_{r_1\cdots r_n})X_1^{r_1}\cdots X_n^{r_n} + u_1\sum r_1 a_{r_1\cdots r_n}X_1^{r_1-1}X_2^{r_2}\cdots X_n^{r_n}\\ &+ \cdots + u_n\sum r_n a_{r_1\cdots r_n}X_1^{r_1}\cdots X_{n-1}^{r_{n-1}}X_n^{r_n-1}.\end{aligned}\tag{9}$$

像前一小节一样, 这引导我们引进下列记号. 首先令

$$f^{D} = \sum D(a_{r_1\cdots r_n})X_1^{r_1}\cdots X_n^{r_n},$$

这是把 D 应用到 f 的系数所得到的多项式. 其次称多项式

$$f_i' = \sum r_i a_{r_1\cdots r_n}X_1^{r_1}\cdots X_{i-1}^{r_{i-1}}X_i^{r_i-1}X_{i+1}^{r_{i+1}}\cdots X_n^{r_n}\tag{10}$$

为 f 关于 X_i 的**偏导式**. 如果把 f 看作系数在 $K[X_1\cdots X_{i-1}, X_{i+1},\cdots,X_n]$ 内的 X_i 的多项式, 那么 f_i' 正是前一小节意义下的 f 的导多项式 (这必定对应多变量函数的偏导数的经典概念, 其中把其余变量看作常量). 引进了这些记号, 关系 (9) 还可以写成

$$D'(f) = f^{D} + \sum_{i=1}^{n} u_i\cdot f_i'.\tag{11}$$

在实际中, 代替记号 f_i', 经常使用分析中常用的记号

$$f_{X_i}', \quad \frac{\partial f}{\partial X_i}.$$

如果令

$$D_i(f) = f_i',$$

那么显然从 $K[X_1,\cdots,X_n]$ 到自身内的映射 D_i 满足下列条件: 这是一个导子, 并且

$$D_i(a) = 0, \quad 如果\ a\in K,$$

$$D_i(X_j) = \begin{cases}0, & i\neq j,\\ 1, & i=j.\end{cases}$$

并且根据定理 2 或把定理 1 应用到环

$$K[X_1\cdots X_{i-1}, X_{i+1},\cdots,X_n]$$

上, D_i 在这个环上是零, 这些性质刻画了 D_i 的特征.

4. 复合函数的导子

"复合函数的求导" 的经典定理当涉及多项式时是下列结果的一个推论:

定理 3 设 L 是一个交换环, K 是 L 的一个子环, D 是在 K 上为零的 L 的导子, 而 f 是系数在 K 内的 n 个未定元的多项式. 则对任意 $u_1, \cdots, u_n \in \mathrm{L}$ 有

$$\mathrm{D}[f(u_1, \cdots, u_n)] = \sum_{i=1}^{n} f_i'(u_1, \cdots, u_n)\mathrm{D}(u_i).$$

显然, 只需当 f 是单项式时确立这个结果. 设

$$f = a\mathrm{X}_1^{r_1} \cdots \mathrm{X}_n^{r_n},$$

我们有

$$\begin{aligned}\mathrm{D}[f(u_1, \cdots, u_n)] &= \mathrm{D}(au_1^{r_1} \cdots u_n^{r_n}) \\ &= \mathrm{D}(a)u_1^{r_1} \cdots u_n^{r_n} + \sum_{i=1}^{n} au_1^{r_1} \cdots u_{i-1}^{r_{i-1}}\mathrm{D}(u_i^{r_i})u_{i+1}^{r_{i+1}} \cdots u_n^{r_n} \\ &= \mathrm{D}(a)u_1^{r_1} \cdots u_n^{r_n} + \sum_{i=1}^{n} \mathrm{D}(u_i)r_i au_1^{r_1} \cdots u_{i-1}^{r_{i-1}}u_i^{r_i-1}u_{i+1}^{r_{i+1}} \cdots u_n^{r_n}.\end{aligned}$$

如果在 K 上 $\mathrm{D} = 0$, 那么第一项消失, 剩下的显然就是所寻求的公式.

推论 设 K 是一个交换环, f 是系数在 K 内的 n 个未定元的多项式, 而

$$u_1, \cdots, u_n \in \mathrm{K}[\mathrm{Y}_1, \cdots, \mathrm{Y}_p]$$

是系数在 K 内的 p 个未定元的多项式. 多项式

$$g(u_1, \cdots, u_n) = f[u_1(\mathrm{Y}_1, \cdots, \mathrm{Y}_p), \cdots, u_n(\mathrm{Y}_1, \cdots, \mathrm{Y}_p)]$$

的偏导式由关系

$$\frac{\partial g}{\partial \mathrm{Y}_j} = \sum_{i=1}^{n} f_i'(u_1, \cdots, u_n) \cdot \frac{\partial u_i}{\partial \mathrm{Y}_j} \quad (1 \leqslant j \leqslant p)$$

给定.

为了确立此式只需应用定理 3, 其中取 $\mathrm{L} = \mathrm{K}[\mathrm{Y}_1, \cdots, \mathrm{Y}_n]$, 取 D 为关于 Y_j 的偏导子.

这个推论是严格意义上的多项式的复合函数定理.

讲述了由本节的定义几乎平凡地得到的这个结果, 不会再期待更加深刻的推论.

5. Taylor 公式

设 K 是一个交换环, f 是系数在 K 内的一个未定元的多项式, 考虑系数在 K 内的两个未定元 X 和 Y 的多项式 X + Y. 把 X + Y 代入 f 中出现的未定元, 我们得到一个多项式

$$f(\mathrm{X}+\mathrm{Y}) \in \mathrm{K}[\mathrm{X},\mathrm{Y}] = \mathrm{K}[\mathrm{X}][\mathrm{Y}],$$

可以把它写成系数在 K[X] 内的 Y 的多项式, 如果 f 是 n 次的, 其形式是

$$f(\mathrm{X}+\mathrm{Y}) = f_0(\mathrm{X}) + f_1(\mathrm{X})\mathrm{Y} + \cdots + f_n(\mathrm{X})\mathrm{Y}^n. \tag{12}$$

我们打算借助 f 的**逐次导式**, 即

$$f'' = (f')', f''' = (f'')', \cdots$$

计算多项式 $f_p(\mathrm{X})$. 为此, 关于 Y 对于关系式 (12) 两端求导. 根据定理 3 的推论, 左端的导式是 $f'(\mathrm{X}+\mathrm{Y})$, 这样就得到

$$f'(\mathrm{X}+\mathrm{Y}) = f_1(\mathrm{X}) + 2f_2(\mathrm{X})\mathrm{Y} + \cdots + nf_n(\mathrm{X})\mathrm{Y}^{n-1}.$$

重新关于 Y 对于这个结果求导, 我们得到

$$f''(\mathrm{X}+\mathrm{Y}) = 2f_2(\mathrm{X}) + 3\cdot 2f_3(\mathrm{X})\mathrm{Y} + \cdots + n(n-1)f_n(\mathrm{X})\mathrm{Y}^{n-2},$$

如此继续, 显然得到

$$f^{(k)}(\mathrm{X}+\mathrm{Y}) = k!f_k(\mathrm{X}) + (k+1)k\cdots 2f_{k+1}(\mathrm{X})\mathrm{Y} + \cdots + n(n-1)\cdots(n-k+1)f_n(\mathrm{X})\mathrm{Y}^{n-k}.$$

显然这个关系是系数在 K 内的 X 和 Y 的多项式之间的一个等式, 如果用 u 和 v 代换其中的 X 和 Y 等式保持成立, 这里 u 和 v 是包含 K 的交换环 L 的任意元素. 在特殊情形, 用 X 和 0 代换 X 和 Y, 得到的是关系

$$f^{(k)}(\mathrm{X}) = k!f_k(\mathrm{X}), \quad 0 \leqslant k \leqslant n.$$

故得

定理 4 设 f 是系数在一个交换环 K 内的一个未定元的 n 次多项式, 而 X 和 Y 是 K 上的两个未定元. 则有

$$f(\mathrm{X}+\mathrm{Y}) = f(\mathrm{X}) + f'(\mathrm{X})\mathrm{Y} + f_2(\mathrm{X})\mathrm{Y}^2 + \cdots + f_n(\mathrm{X})\mathrm{Y}^n,$$

其中

$$k!f_k(\mathrm{X}) = f^{(k)}(\mathrm{X}) \quad 对于\ 2 \leqslant k \leqslant n. \tag{13}$$

这个公式称为 **Taylor 公式**的理由如下. 首先由于这是多项式之间的一个等式, 因此可以在其中用任意一个包含 K 的交换环的任意元素代换 X 和 Y, 在特殊情形可以用 K 的 x 和 h 代换 X 和 Y, 于是对所有 $x, h \in \mathrm{K}$, 有

$$f(x+h) = f(x) + f'(x)h + f_2(x)h^2 + \cdots + f_n(x)h^n, \quad \text{其中 } k!f_k(x) = f^{(k)}(x),$$

现在假定 K = 实数域 $\mathbf{R}$, 则显然有

$$f_k(x) = \frac{f^{(k)}(x)}{k!},$$

公式变成

$$f(x+h) = f(x) + h\frac{f'(x)}{1!} + h^2\frac{f''(x)}{2!} + \cdots + h^n\frac{f^{(n)}(x)}{n!},$$

这是古典的 Taylor 公式 (这个公式在多项式函数情形是纯代数性质的一个结果).

例 2 取 $\mathrm{K} = \mathbf{Z}$ 和

$$f(\mathrm{X}) = \mathrm{X}^n;$$

公式 (12) 和 (13) 写成

$$(\mathrm{X}+\mathrm{Y})^n = \mathrm{X}^n + n\mathrm{X}^{n-1}\mathrm{Y} + f_2(\mathrm{X})\mathrm{Y}^2 + \cdots + f_n(\mathrm{X})\mathrm{Y}^n,$$

其中

$$k!f_k(\mathrm{X}) = n(n-1)\cdots(n-k+1)\mathrm{Y}^{n-k},$$

或写成

$$k![f_k(\mathrm{X}) - \binom{n}{k}\mathrm{X}^{n-k}] = 0.$$

由于 $\mathbf{Z}$ 是整环, 由此得到

$$f_k(\mathrm{X}) = \binom{n}{k}\mathrm{X}^{n-k},$$

于是在这种情形得到关系

$$(\mathrm{X}+\mathrm{Y})^n = \sum_{k=0}^{n}\binom{n}{k}\mathrm{X}^{n-k}\mathrm{Y}^k.$$

为了推出 §8 第 4 小节所建立的二项式公式, 只需注意到如果带有理整数系数的两个多项式 $f(\mathrm{X},\mathrm{Y})$ 和 $g(\mathrm{X},\mathrm{Y})$ 相等, 则关系

$$f(x,y) = g(x,y)$$

当 x 和 y 是任意交换环 L 的任意元素时都成立.

6. 交换域的特征

回到定理 4. 要使公式 (12) 真正有益处, 必须能够完全计算多项式 f_k, 而为此不得不试图从公式 (13) 求出这些多项式. 换句话说, 给定有理整数 $r \neq 0$ (在当前情形是 $k!$) 和 K 的一个元素 b (在当前情形, 多项式 $f^{(k)}(\mathrm{X})$ 的任意一个系数), 求所有的 $x \in \mathrm{K}$, 使得

$$r \cdot x = b.$$

因为给定

$$rx = (r \cdot 1)x,$$

其中 1 是 K 的单位元, 所以一旦 $r \cdot 1$ 是可逆的问题就有唯一解. 如果 K 是一个交换域, 这表明 $r \cdot 1 \neq 0$.

在 K 是一个交换域情形下, 我们被引导至在有理整数环 **Z** 里, 考虑使得

$$r \cdot 1 = 0$$

的整数 r 的集合 I. I 含有 0, 如果它含有 r 和 s, 则它显然含有 $r - s$, 故这是加法群 **Z** 的一个子群, 因此 (§7, 例 8) 存在唯一的一个整数 $p \geqslant 0$, 使得 $\mathrm{I} = p\mathbf{Z}$, 我们说 p 是**域 K 的特征**.

一个交换域 K 的特征 p 显然可以如下定义: 如果 $r \cdot 1 = 0$ 蕴含 $r = 0$, 那么 $p = 0$, 否则 p 是使得 $p \cdot 1 = 0$ 的最小正整数. 在所有情形, 关系 $r \cdot 1 = 0$ 表示 r 是 p 的倍数, 或还有: $r \cdot 1 \neq 0$, 必须并且只需 r 不能被 K 的特征整除.

重要的是要注意一个交换域的特征或者是 0, 或者是一个素数. 事实上, 假定域 K 的特征 $p \neq 0$, 考虑两个这样的正整数 r 和 s, 它们使得

$$p = rs,$$

于是有

$$0 = p \cdot 1 = (r \cdot 1)(s \cdot 1).$$

由于域是一个整环, 由此得到或者 $r \cdot 1 = 0$ (这时 p 整除 r) 或者 $s \cdot 1 = 0$ (这时 p 整除 s), 这就确立了我们的断言.

如果 K 是一个特征为 0 的域, 则对于所有 $b \in \mathrm{K}$ 和所有整数 $r \neq 0$, 方程

$$rx = b$$

在 K 内具有唯一解, 即

$$x = (r \cdot 1)^{-1}b,$$

更简单地可以写成形式

$$x = \frac{b}{r} \quad 或 \quad b/r.$$

例 3 显然域 $\mathbf{Q}, \mathbf{R}, \mathbf{C}$ 以及更一般的所有包含 $\mathbf{Q}$ 的交换域的特征是 0.

例 4 取 $\mathrm{K} = \mathbf{Z}/p\mathbf{Z}$, 这里 p 是素数 (§8, 定理 1). 对于所有有理整数 r, 在 $\mathbf{Z}/p\mathbf{Z}$ 里关系 $r \cdot 1 = 0$ 显然意味着 r 作为模 p 整数是 0, 即 r 是 p 的倍数. 故这里有 $\mathrm{I} = p\mathbf{Z}$, 即对于素数 p, 域 $\mathbf{Z}/p\mathbf{Z}$ 的特征是 p.

注 1 设 K 是特征为 0 的域, 从 $\mathbf{Z}$ 到 K 的映射



$$n \to n \cdot 1$$

是单射的, 因此是从 $\mathbf{Z}$ 到 K 的一个子环上的同构. 事实上, 这个映射可以延拓为一个从域 $\mathbf{Q}$ 到 K 的一个子域上的同构. 事实上, 设 $x \in \mathbf{Q}$, 并且写出 $x = a/b$, 其中 $a, b \in \mathbf{Z}, b \neq 0$, 那么 K 的元素

$$j(x) = (a \cdot 1)(b \cdot 1)^{-1}$$

仅依赖 x, 而不依赖其分数表示. 事实上, 关系 $a'/b' = a''/b''$ 写成 $a'b'' = a''b'$, 故蕴含

$$(a' \cdot 1)(b'' \cdot 1) = (a'' \cdot 1)(b' \cdot 1),$$

于是有

$$(a' \cdot 1)(b' \cdot 1)^{-1} = (a'' \cdot 1)(b'' \cdot 1)^{-1},$$

这就验证了我们的断言. j 定义之后, 那么就可以直接验证如此得到的从 $\mathbf{Q}$ 到 K 内的映射 j 是从 $\mathbf{Q}$ 到 K 的一个子域上的同构.

在实际中, 经常把每个有理数 $x \in \mathbf{Q}$ 等同于它在 K 内的像 $j(x)$, 于是 $\mathbf{Q}$ 是所有特征为 0 的域的一个子域.

还可以证明 (参见习题 8) 所有特征为 $p \neq 0$ 的域 K 包含一个同构于 $\mathbf{Z}/p\mathbf{Z}$ 的子域, 即 K 的单位元的整倍数 $r \cdot 1 (r \in \mathbf{Z})$ 的集合.

设 K 是一个以 0 为特征的域, 那么显然前一个小节的 Taylor 公式取形式

$$f(\mathrm{X}+\mathrm{Y}) = \sum_{k=0}^{n} f^{(k)}(\mathrm{X}) \cdot \frac{\mathrm{Y}^k}{k!}. \tag{14}$$

这个结果对于特征 $p \neq 0$ 也成立, 只要对于 $k \leqslant n$ 有 $k! \cdot 1 \neq 0$, 即 p 不整除任何小于或等于 n 的整数, 也就是说只要

$$n < p.$$

7. 方程根的重数

设 K 是一个交换环, f 是系数在 K 内的一个未定元的多项式, 在 Taylor 公式

$$f(\mathrm{X}+\mathrm{Y}) = f(\mathrm{X}) + f'(\mathrm{X})\mathrm{Y} + f_2(\mathrm{X})\mathrm{Y}^2 + \cdots$$

里用 a 和 $\mathrm{T}-a$ 代换 X 和 Y, 这里 a 是 K 的一个给定的元素, 而 T 则是 K 上的一个未定元. 此时有

$$\begin{aligned} f(\mathrm{T}) &= f(a) + f'(a)(\mathrm{T}-a) + f_2(a)(\mathrm{T}-a)^2 + \cdots \\ &= f(a) + (\mathrm{T}-a)q(\mathrm{T}), \end{aligned}$$

其中 $q \in \mathrm{K}[\mathrm{T}]$. 给定系数在 K 内的两个多项式 g, h, 说 h 被 g **整除**, 如果存在第三个系数在 K 内的多项式 q, 使得 $h = gq$. 从我们刚得到的结果直接推出如下定理 (§28, 引理 1 已经确立):

定理 5 设 f 是系数在交换环 K 内的一个未定元的多项式. K 的一个元素 a 是 f 的根, 必须并且只需多项式 $f(\mathrm{X})$ 被多项式 $\mathrm{X}-a$ 整除.

事实上, 如果 a 是 f 的根, 则有 $f(a)=0$, 那么 Taylor 公式给出

$$f(\mathrm{T}) = (\mathrm{T}-a)q(\mathrm{T}).$$

反之, 由这个关系得到

$$f(a) = (a-a)q(a) = 0,$$

定理证毕.

给定多项式 f 的一个根 a, 使得多项式 $f(\mathrm{T})$ 被多项式

$$(\mathrm{T}-a)^r$$

整除的最大整数 r 称为 a **的重数**. 如果重数 $r=1$, 则说 a 是**单根**, 如果重数 $r=2$, 则说 a 是**二重根**, 等等.

定理 6 设 f 是系数在交换环 K 内的一个未定元的多项式. 元素 $a \in \mathrm{K}$ 是 f 的单根, 必须并且只需

$$f(a) = 0, \quad f'(a) \neq 0.$$

事实上, 如果 a 是根, 则有 $f(\mathrm{T}) = (\mathrm{T}-a)q(\mathrm{T})$, 其中

$$q(\mathrm{T}) = f'(a) + f_2(a)(\mathrm{T}-a) + \cdots,$$

故

$$q(a) = f'(a).$$

假定 $f'(a)=0$, 那么 (定理 5) 多项式 $q(\mathrm{T})$ 被 $\mathrm{T}-a$ 整除, 因此 f 被 $(\mathrm{T}-a)^2$ 整除, a 不是单根. 反之, a 不是单根, 则有

$$f(\mathrm{T}) = (\mathrm{T}-a)^2 g(\mathrm{T}),$$

由此得到

$$f'(\mathrm{T}) = 2(\mathrm{T}-a)g(\mathrm{T}) + (\mathrm{T}-a)^2 g'(\mathrm{T}),$$

这显然保证 $f'(a)=0$, 并且完成了证明.

定理 7 设 f 是系数在 0 特征的交换域 K 内的一个未定元的多项式. K 的一个元素 a 是 f 的 r 重根, 必须并且只需以下关系成立:

$$f(a) = f'(a) = \cdots = f^{(r-1)}(a) = 0, \quad f^{(r)}(a) \neq 0. \tag{15}$$

假定 a 是 r 重根, 则有

$$f(\mathrm{T}) = (\mathrm{T}-a)^r g(\mathrm{T}), \tag{16}$$

其中 $g(a)\neq 0$, 否则 (定理 5) $g(\mathrm{T})$ 将是被 $\mathrm{T}-a$ 整除的, 因此 $f(\mathrm{T})$ 是被 $(\mathrm{T}-a)^{r+1}$ 整除的. 对于关系 (16) 求导 k 次显然得到关系

$$f^{(k)}(\mathrm{T}) = r(r-1)\cdots(r-k+1)(\mathrm{T}-a)^{r-k}g(\mathrm{T}) + (\mathrm{T}-a)^{r-k+1}q_k(\mathrm{T}),$$

其中 $q_k(\mathrm{T})$ 是一个精确形式无关紧要的多项式. 从这个关系得到对于 $k\leqslant r-1$ 有 $f^{(k)}(a)=0$, 此外有

$$f^{(r)}(a) = r!g(a).$$

由于 $g(a)\neq 0$, 并且 K 是 0 特征的, 必然有 $f^{(r)}(a)\neq 0$.

反之, 假定条件 (15) 成立, 由于 K 是 0 特征的, 可以写出

$$f(\mathrm{T}) = \sum_{k=0}^{n}(\mathrm{T}-a)^k\frac{f^{(k)}(a)}{k!} = (\mathrm{T}-a)^r\left[\frac{f^{(r)}(a)}{r!} + (\mathrm{T}-a)\frac{f^{(k+1)}(a)}{(k+1)!} + \cdots\right],$$

即有关系

$$f(\mathrm{T}) = (\mathrm{T}-a)^r g(\mathrm{T}),$$

其中 $g(a)\neq 0$. 如果 f 是被 $(\mathrm{T}-a)^{r+1}$ 整除的, 由于 K[T] 是整环, 那么显然 g 将是被 $\mathrm{T}-a$ 整除的, 由于 $g(a)$ 不是零, 这是不可能的. 因此 $(\mathrm{T}-a)^r$ 就是整除 f 的 $\mathrm{T}-a$ 的最高次数幂, 故 a 是 r 重根, 这就完成了证明.

例 5 K 是一个 0 特征域, 探索在什么条件下系数 p, q 在 K 内的方程

$$x^3 + px + q = 0$$

具有重根 a. 这个重根应当使得方程左端及其导式为零 (定理 6), 即满足

$$a^3 + pa + q = 0,$$
$$3a^2 + p = 0.$$

第一个关系乘以 3, 第二个乘以 a, 两端分别相减即得 $2pa+3q=0$, 即

$$a=-3q/2p \quad (p\neq 0);$$

把这个结果代入关系 $3a^2+p=0$, 即得系数 p 和 q 应当满足

$$4p^3+27q^2=0. \tag{17}$$

反之, 如果这个关系成立, 就可直接发现 $-3q/2p(p\neq 0)$ 是所考虑的方程的重根. 当 $p=0$ 时, 只有在 $q=0$ 时有重根因此条件 (17) 是有重根的充分和必要条件.

读者可以验证, 事实上仅假定 K 的特征异于 2 和 3, 而无需假定 K 的特征是 0, 这个结果也成立.

§30 习题

1. 设 K 是一个特征为 0 的交换域 (例如 $\mathrm{K}=\mathbf{C}$). 证明方程

$$\frac{x^n}{n!}+\frac{x^{n-1}}{(n-1)!}+\cdots+\frac{x}{1!}+1=0$$

没有任何重根.

2. 设 K 是一个交换域. 求一个 7 次多项式 $f\in\mathrm{K[X]}$, 使得 1 至少是 $f(\mathrm{X})+1$ 的 4 重根, 而 -1 至少是 $f(\mathrm{X})-1$ 的 4 重根. 用 n 和 $2n-1$ 代换 4 和 7 推广之.

3. 下列多项式中的每一个都有根 1, 确定其重数:

$$\mathrm{X}^{2n}-n\mathrm{X}^{n+1}+n\mathrm{X}^{n-1}-1;$$
$$\mathrm{X}^{2n+1}-(2n+1)\mathrm{X}^{n+1}+(2n+1)\mathrm{X}^{n}-1;$$
$$\mathrm{X}^{2n}-n^2\mathrm{X}^{n+1}+2(n^2-1)\mathrm{X}^{n}-n^2\mathrm{X}^{n-1}+1.$$

4. 设 f 是其系数在一个交换域 K 内的一个未定元的多项式. 假定 $f'=0$. 证明 f 是常多项式, 如果 K 是特征为 0 的; f 是 X^p 的多项式, 如果 K 是特征为 $p\neq 0$ 的. 其逆如何?

¶5. 设 K 是特征为 0 的一个交换域. 我们打算求所有的从 K 到矩阵环 $\mathrm{M}_n(\mathrm{K})$ 内的映射

$$t\to \mathrm{U}(t),$$

它们要满足

$$\mathrm{U}(x+y)=\mathrm{U}(x)\mathrm{U}(y) \quad \text{对于任意 } x,y\in\mathrm{K},$$
$$\mathrm{U}(0)=1_n,$$

并且它们是多项式函数, 即

$$\mathrm{U}(t)=\mathrm{A}_0+\mathrm{A}_1t+\cdots+\mathrm{A}_rt^r+\cdots$$

其中的矩阵 $\mathrm{A}_r\in\mathrm{M}_n(\mathrm{K})$ 几乎全部为零.

a) 证明多项式函数 $\mathrm{U}(t)$ 的导式满足

$$\mathrm{U}'(t) = \mathrm{A}_1 \cdot \mathrm{U}(t).$$

b) 证明矩阵 $\mathrm{A}_1 = \mathrm{N}$ 是幂零的, 并且

$$\mathrm{U}(t) = \sum_{r \geqslant 0} \mathrm{N}^r \frac{t^r}{r!} = \exp(t\mathrm{N})$$

(参见 §8, 习题 2).

c) 证明, 反之, 对于所有幂零矩阵 N, 映射

$$t \to \exp(t\mathrm{N})$$

满足所需要的条件.

d) 求在 §12 习题 11 的函数 $\mathrm{U}(t)$ 的情形的矩阵 N, 并且验证在这个情形有 $\mathrm{U}(t) = \exp(t\mathbf{N})$.

6. 设 K 是特征为 0 的一个交换域. 证明不存在任何非零多项式 $f \in \mathrm{K}[\mathrm{X}]$, 使得对于任何 $x, y \in \mathrm{K}$ 有

$$f(x+y) = f(x)f(y).$$

对于关系

$$f(xy) = f(x) + f(y)$$

回答同样的问题. 满足关系

$$f(x+y) = f(x) + f(y)$$

的多项式是什么?

¶7. 设 K 是特征 $p \neq 0$ 的域.

a) 证明对于任意 $x, y \in \mathrm{K}$ 我们有

$$(x+y)^p = x^p + y^p.$$

由此推出如果 q 是 p 的一个幂, 则有

$$(x+y)^q = x^q + y^q.$$

b) 证明 $x \to x^p$ 是从 K 到 K 的一个子域 (记为 K^p) 上的同构. 证明如果 K 是有限的, 则 $\mathrm{K}^p = \mathrm{K}$.

c) 对于任意 $x, y \in \mathrm{K}$ 满足

$$f(x+y) = f(x) + f(y)$$

的 $f \in \mathrm{K}[\mathrm{X}]$ 是什么?

¶8. 设 K 是特征 $p \neq 0$ 的域. 证明对于 $n \in \mathbf{Z}$ 和 $x \in \mathrm{K}$, 元素 nx 仅依赖于 x 和 n 模 p 的类. 由此推出可以考虑 K 为域 $\mathbf{Z}/p\mathbf{Z}$ 上的向量空间.

证明特征为 p 的有限域的元素个数是 p 的一个幂.

9. 设 p 是一个素数. 证明二项式系数 $\binom{p^n}{r}$ 对于 $n \geqslant 1$ 和所有满足 $1 \leqslant r \leqslant p^n - 1$ 的 r 是被 p 整除的. (对于域 $\mathrm{K} = \mathbf{Z}/p\mathbf{Z}$ 利用习题 7.) 能给一个初等的证明吗?

10. 设 $f(\mathrm{X}_1, \cdots, \mathrm{X}_n)$ 是系数在一个交换环 K 内的 n 个未定元的多项式. 假定 f 是 r 次齐次的. 证明

$$\mathrm{X}_1 f_1'(\mathrm{X}_1, \cdots, \mathrm{X}_n) + \cdots + \mathrm{X}_n f_n'(\mathrm{X}_1, \cdots, \mathrm{X}_n) = r \cdot f(\mathrm{X}_1, \cdots, \mathrm{X}_n),$$

其中 f_i' 是 f 关于 X_i 的偏导多项式. 这个关系 (称为 Euler 等式) 刻画了 r 次齐次多项式的特征吗?

11. 设

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 = 0 \tag{$*$}$$

是系数 a_i 为有理整数的代数方程, 下面假定 a_i 是互素的 (通过除以 a_i 的最大公约数总可以化成这种情形). 设 $x = p/q$ 是方程 $(*)$ 的一个根, 假定 p 和 q 是互素的. 证明 p 整除 a_0, q 整除 a_n.

应用: 求下列方程的有理根:

$$6x^4 - 11x^3 - x^2 - 4 = 0,$$
$$2x^3 + 12x^2 + 13x + 15 = 0,$$
$$6x^5 + 11x^4 - x^3 + 5x - 6 = 0,$$
$$x^6 + 3x^5 + 4x^4 + 3x^3 - 15x^2 - 16x + 20 = 0,$$
$$2x^6 + x^5 - 9x^4 - 6x^3 - 5x^2 - 7x + 6 = 0.$$

12. 设 K 是一个交换环. 考虑由 K 和一个满足关系

$$\varepsilon^2 = 0$$

(在 §§27, 28 的习题 23 中令 $n = 1$) 的元素 ε 生成的环 $\mathrm{L} = \mathrm{K}[\varepsilon]$. 从 K 到 L 的映射

$$x \to x + \mathrm{D}(x)\varepsilon$$

是一个同态, 必须并且只需 D 是环 K 的一个导子.

13. 在怎样的域内有等式

$$x^4 - x^2 + 1 = (x^2 - 5x + 1)(x^2 + 5x + 1)?$$

¶14. 设 K 是一个特征 $p \neq 0$ 的域. 如果 $x \in \mathrm{K}$ 对于一个整数 n 满足 $x^n = 1$, 则存在一个不被 p 整除的 r, 使得

$$x^r = 1.$$

¶15. 设 K 是一个交换环. (用关于整数 $n \geqslant 0$ 的归纳法) 定义在 K 上的至多 n 次的微分运算: 这是从 K 到 K 内的映射 D, 它满足

$$\mathrm{D}(x + y) = \mathrm{D}(x) + \mathrm{D}(y) \quad 对于任意\ x, y \in \mathrm{K},$$

并且还具有下列性质: 如果 $n = 0$, 那么存在一个 $a \in \mathrm{K}$, 使得

$$\mathrm{D}(x) = ax \quad 对于所有\ a \in \mathrm{K};$$

如果 $n \geqslant 1$, 则对于所有 $x \in \mathrm{K}$, 存在 K 上的一个至多 $n-1$ 阶的微分算子 D_x, 使得

$$\mathrm{D}(xy) = x \cdot \mathrm{D}(y) + \mathrm{D}_x(y) \quad \text{对于所有 } y \in \mathrm{K}.$$

a) 确定 K 上的所有至多 1 阶的微分算子.

b) 证明如果 D' 和 D'' 分别是至多 r 和 s 阶的微分算子, 则复合映射

$$\mathrm{D}'' \circ \mathrm{D}'$$

是一个至多 $r+s$ 的微分算子, 并且 “Jacobi 括号”

$$\mathrm{D}'' \circ \mathrm{D}' - \mathrm{D}' \circ \mathrm{D}''$$

是至多 $r+s-1$ 阶微分算子.

c) 设 D 是一个至多 n 阶微分算子. 考虑 K 的一族元素 $(x_i)_{i \in \mathrm{I}}$, 其元素个数 $\mathrm{Card}(\mathrm{I}) = n+1$. 对于 I 的所有子集 F, 令

$$x_{\mathrm{F}} = \prod_{i \in \mathrm{F}} x_i \quad \text{和} \quad x_{\varnothing} = 1.$$

证明等式

$$\sum_{\mathrm{F} \subset \mathrm{I}} (-1)^{\mathrm{Card}(\mathrm{F})} x_{\mathrm{F}} D(x_{\mathrm{I}-\mathrm{F}}) = 0.$$

证明, 反之, 所有从 K 到 K 内的映射 D, 如果具有这个性质, 并且满足关系 $\mathrm{D}(x+y) = \mathrm{D}(x) + \mathrm{D}(y)$, 则 D 是 K 上的一个至多 n 阶微分算子.

d) 由上面的结果通过关于 n 的归纳法推出计算 $n+1$ 个多项式的乘积的 p 阶偏导式的公式. 在 $p=1$ 这种情形公式如何?

e) 取

$$\mathrm{K} = k[\mathrm{X}_1, \cdots, \mathrm{X}_r],$$

其中的 K 是一个交换环. 构造所有在 k 上为零的微分算子.

§31 主理想整环

我们提醒 (§8, 例 10) 称其所有理想皆为主理想的整环为*主理想整环*. 下一节将看到如果 K 是一个域, 则系数在 K 内的一个未定元的多项式环 K[X] 是一个主理想整环. 在这一节我们要建立主理想整环的若干算术性质, 这些性质推广了整数的相应性质, 并且将在下一节应用到多项式. 在整个这一节 K *表示一个主理想整环*.

1. 最大公因子

设 K 是一个主理想整环, 而 $x_1, \cdots, x_n$ 是 K 的元素. 设

$$\mathrm{I} = (x_1, \cdots, x_n)$$

是由 $x_1, \cdots, x_n$ 生成的 K 的理想[(*)]. 由于 K 是主理想整环, 这个理想由一个元素 d 生成, 不计下述差别 d 是唯一的, d 可以代以 ud, u 是环 K 的任意可逆元[(**)].

称使得

$$(x_1, \cdots, x_n) = (d) \tag{1}$$

成立的 K 的所有元素 d 为 $x_1, \cdots, x_n$ 的**最大公因子**. 由于左端是这样的 $y \in \mathrm{K}$ 的集合, 存在 $u_1, \cdots, u_n \in \mathrm{K}$, 使得

$$y = u_1 x_1 + \cdots + u_n x_n, \tag{2}$$

我们发现还可以用下列事实刻画 d, 形式为 (2) 的元素的集合等于 d 在 K 内的倍元的集合. 我们考虑一个特殊情形, 由于 $d \in (d)$, 我们发现存在 $u_1, \cdots, u_n \in \mathrm{K}$, 使得

$$d = u_1 x_1 + \cdots + u_n x_n, \tag{3}$$

这个结果以 Bezout 定理命名.

为了表明 d 的意义而采用 "最大公因子" 的术语由以下结果证实其合理性: K 的一个元素同时整除 $x_1, \cdots, x_n$, 必须并且只需它整除 d.

事实上关系 (1) 表明理想 (d) 含有各个 x_i, 因此后者是 d 的倍元, 因此显然 d 的所有因子整除 x_i. 反之, 设 $m \in \mathrm{K}$ 是 $x_1, \cdots, x_n$ 的一个公因子, 写出

$$x_i = m y_i \quad (1 \leqslant i \leqslant n),$$

代入 (3) 即得

$$d = m(u_1 y_1 + \cdots + u_n y_n),$$

这就验证了 m 整除 d, 并且完成了证明.

例 1 取 **Z** 作为 K, 它是主理想整环 (§10, 例 9). 给定整数 $x_1, \cdots, x_n$, 那么存在一种 "自然的" 或 "典范的" 方式选择理想 $(x_1, \cdots, x_n)$ 的一个生成元, 即取这个理想的正的生成元. 因此可以像在初等数论中通常的那样谈论 $x_1, \cdots, x_n$ 的唯一的最大公约数. $x_1, \cdots, x_n$ 的公约数也是这个最大公约数的约数, 那么显然最大公约数 (选择正的) 是 $x_1, \cdots, x_n$ 的所有公约数中的最大者. 我们重新回到经典的概念, 并且还补充了 Bezout 定理 (在初等数论中是不讲的). 参见 §7 的例 8.

(*) 希望初学的读者不要把理想 $(x_1, \cdots, x_n)$ 跟用同样记号表示的 K^n 的元素相混淆! 这里采用的表示由 $x_1, \cdots, x_n$ 生成的理想的记号在数论里是传统的.

(**) 事实上, 假定 $(d) = (d')$, 那么存在 $u, v \in \mathrm{K}$, 使得 $d' = ud, d = vd'$, 由此得到 $vud = d$, 如果 $d \neq 0$, 那么 (由于 K 是整环) 得到 $vu = 1$, 于是 u 在 K 内是可逆的.

2. 互素元素

称 K 的元素 $x_1, \cdots, x_n$ 是**互素的**, 如果它们有一个最大公因子 1.

定理 1 设 $x_1, \cdots, x_n$ 是主理想整环 K 的元素. 则以下性质是等价的:

a) $x_1, \cdots, x_n$ 是互素的;

b) $x_1, \cdots, x_n$ 的所有的公因子是 K 的可逆元;

c) 存在 K 的元素 $u_1, \cdots, u_n \in \mathrm{K}$, 使得

$$u_1x_1 + \cdots + u_nx_n = 1;$$

d) 对于所有的 $y \in \mathrm{K}$, 存在 K 的元素 $u_1, \cdots, u_n \in \mathrm{K}$, 使得

$$y = u_1x_1 + \cdots + u_nx_n.$$

a) 表明

$$(x_1, \cdots, x_n) = (1) = \mathrm{K},$$

由此得 a) 与 d) 等价. 如果 x_i 是互素的, 则它们的公因子是 1 的因子, 即 K 的可逆元; 反之, 如果 x_i 的所有公因子是可逆的, 那么 x_i 的最大公因子 (或更正确地说是之一) 是可逆的, 因为一个可逆元的倍元组成整个环 K, 故有 $(x_1, \cdots, x_n) = \mathrm{K}$, 故性质 a) 和 b) 是等价的. 最后由于 Bezout 定理, a) 蕴含 c). 而 c) 蕴含 d), 这是因为 c) 表明理想 $(x_1, \cdots, x_n)$ 含有 1, 从而是整个 K. 这就完成了证明.

定理 1 的性质 c) 对于证明某些 "经典的" 但初看并非显然的性质十分有用.

定理 2 设 x 和 y 是 K 的两个非零元素, 并且 d 是乘积 xy 的一个因子. 如果 d 和 x 互素, 则 d 整除 y.

由于 d 和 x 互素, 存在 $u, v \in \mathrm{K}$, 使得

$$ud + vx = 1.$$

所得的结果乘以 y 即得

$$y = yud + vxy.$$

由于 d 整除 xy 和 yud, 显然整除右端, 故整除 y.

3. 最小公倍

设 $x_1, \cdots, x_n$ 是 K 的非零元素. x_i 的倍元是理想 (x_i) 的元素, 因此 x_i 的公倍是理想 $(x_1) \cap \cdots \cap (x_n)$ 的元素. 这个理想既然是主理想, 就可以提出下列定义: 称使得

$$(x_1) \cap \cdots \cap (x_n) = (m) \tag{4}$$

成立的 K 所有元素 m 为 $x_1, \cdots, x_n$ 的最小公倍. 元素不计一个可逆元是唯一的 (即 x_i 的最小公倍通过 m 乘以 K 的任意可逆元而得到), 而关系 (4) 表明 x_i 的公倍就是 m 的倍元.

定理 3　设 x 和 y 是 K 的两个非零元素. 则有

$$xy = md,$$

其中 m 是 x 和 y 的一个最小公倍, 而 d 是 x 和 y 的一个最大公因子.



注 1　由于有权用 ud 代换 d, 用 vm 代换 m, 这里 u 和 v 是 K 的任意可逆元, 故显然关系

$$xy = md$$

仅当 m 和 d 适当选择时才成立; 正是在这种形式下可以解释定理的表述.

为了证明定理 3, 令

$$x = x'd, \quad y = y'd,$$

并且设 m' 是 x' 和 y' 的最小公倍. 为了 K 的一个元素 z 是 x 的倍元和 y 的倍元, 显然必须并且只需可以写出

$$z = z'd,$$

其中的 z' 是 x' 和 y' 的公倍, 即 m' 的倍元. 如此看来, x 和 y 的公倍是 $m'd$ 的倍元, 故 $m'd$ 是 x 和 y 的最小公倍. 要建立的关系于是写成 $xy = m'd \cdot d$, 约去 d^2 则得

$$x'y' = m'.$$

于是定理 3 将是以下两个引理的推论:

引理 1　设 x 和 y 是 K 的两个非零元素, 而 d 是 x 和 y 是最大公因子. 令 $x = x'd, y = y'd$, 那么 x' 和 y' 是互素的.

事实上, 存在 $u, v \in$ K, 使得 $ux + vy = d$, 约去 d 即得 $ux' + vy' = 1$, 由定理 1 即得引理.

引理 2　如果 x 和 y 是互素的, 则 xy 是 x 和 y 的最小公倍.

设 m 是 x 和 y 的一个公倍. 令 $m = xz$, 则元素 y 整除 xz. 由于 y 与 x 互素, y 整除 z (定理 2), 由此即得引理.

引理 2 可以推广如下:

定理 4　如果 $x_1, \cdots, x_n$ 是 K 的非零元素, 且两两互素, 则 $x_1 \cdots x_n$ 是 $x_1, \cdots, x_n$ 的最小公倍.

对于 $n=2$ 这是引理 2. 我们指出如果定理对于 $n-1$ 个因子的乘积成立, 则对于 n 个因子的乘积也成立.

先用关于 n 的归纳法指出 x_n 是与乘积 $x_1\cdots x_{n-1}$ 互素的. 如果 $n=2$ 这是 x_i 两两互素的假设. 假设已经确认 x_n 是与乘积 $x_1\cdots x_{n-2}$ 互素的, 并且设 d 是 x_n 与 $x_1\cdots x_{n-1}$ 的公因子. 由于 d 整除 x_n, x_n 是与 x_{n-1} 互素的, 故显然 d 与 x_{n-1} 是互素的. 由于 d 整除 $x_1\cdots x_{n-1}$, 故 d 整除 $x_1\cdots x_{n-2}$ (定理 2), 但是根据归纳假设, x_n 和 $x_1\cdots x_{n-2}$ 互素, 因此 d 是可逆的, 由此得到我们的断言.

再用关于 n 的归纳法证明定理 4. 设 m 是 x_i 的公倍. 已经假定定理对于 $n-1$ 个因子的乘积成立, 我们看到 m 是 $x_1\cdots x_{n-1}$ 的一个倍元, 也是 x_n 的倍元, K 的这两个元素我们刚看到是互素的, 故根据引理 2, m 是它们的乘积的一个倍元, 这就完成了证明.

4. 素因子的存在性

称 K 的一个元素是**素元**, 或**不可约元**, 或**极端元**, 如果它不是可逆的, 并且它的仅有的因子是显然的因子, 即 K 的可逆元以及元素 pu, 其中的 u 是可逆的.

当 $\mathrm{K}=\mathbf{Z}$ 时, 显然回到经典的素数概念. 而我们知道在这一情形, 所有 $x\in\mathrm{K}$ 可以写成素数乘积的形式. 我们要把这个结果推广到主理想整环:

定理 5 设 K 是一个主理想整环, 则 K 的所有非零元素是 K 的一个可逆元和若干个素元的乘积.

或者表述为: 所有既非零又非可逆的元素 $x\in\mathrm{K}$ 是若干个素元的乘积 (一个可逆元显然不能分解成素元的乘积, 因为可逆元的因子是可逆元, 从而绝不是素元).

为了证明定理 5, 首先注意到一个主理想整环更是一个 Noether 环: 它的所有理想显然是有限生成的. 因此 (§18, 定理 4) 有下列结果:

引理 3 设 X 是主理想整环 K 的理想的一个非空集合, 那么 X 至少具有一个极大元, 即存在一个 $\mathrm{I}\in\mathrm{X}$, 它不包含于任何另外的 $\mathrm{J}\in\mathrm{X}$ 内.

回忆了这个结果, 我们要用归谬推理证明定理 5. 设 X 是 K 的理想 $\mathrm{I}=(x)$ 的集合, 其中的 x 是 K 的非零元并且不能够写成 K 的一个可逆元和若干个素元的乘积. 为了证明定理, 一切都归结为证明 X 是空集. 如果不然, 它至少含有一个极大元, 设为 (a). 元素 a 既然不能够写成 K 的一个可逆元和素元的乘积, 它自己就既不是可逆元, 也不是素元. 因此可以写出 $a=bc$, 其中的 b 和 c 都不是可逆的, 那么显然 (b) 和 (c) 严格包含 (a); 由于 (a) 不严格包含于任何属于 X 的理想内, 故理想 (b) 和 (c) 都不属于集合 X, 于是定理 5 将对于 b 和 c 成立. 但那样将对于 a 也成立, 这与假设 $(a)\in\mathrm{X}$ 相抵触, 这就结束了证明.

注 2 当 $\mathrm{K}=\mathbf{Z}$ 时自然可以绕开引理 3 (认真说来, 这种情形的初等证明虽



然不明显地引用引理 3, 不过说到底还是遗憾地用到了这个引理; 经典的证明在于: 如果不能分解成素因子的整数存在, 引进这样的整数中的最小者 $a > 0$, 然后注意到 a 不可能是素数, 故可以写出 $a = bc$, 这里 $0 < b < a$, 并且 $0 < c < a$, 在这种情形 b 和 c 可以分解为素因子的乘积, 从而 a 也如此. 显然一般情形的证明直接模仿了这个传统推理). 在下一节会看到当 $\mathrm{K} = \mathrm{L}[\mathrm{X}]$ 且 L 是一个域时, 也可以绕过引理 3 (或者宁肯顺便给出简单的证明).

我们提醒读者, 无论如何, 引理 3 的证明确实是十分简单的, 可以如下进行. 如果引理 3 不成立, 就可以从任意一个理想 $\mathrm{I}_1 \in \mathrm{X}$ 出发, 构造一个严格包含 I_1 的 $\mathrm{I}_2 \in \mathrm{X}$, 再构造一个严格包含 I_2 的 $\mathrm{I}_3 \in \mathrm{X}$, 如此无限继续下去. 为了得到一个矛盾, 一切都归结为指出 K 的理想的整个递增序列

$$\mathrm{I}_1 \subset \mathrm{I}_2 \subset \cdots \subset \mathrm{I}_n \subset \cdots$$

是稳定的, 即存在一个整数 r, 使得

$$\mathrm{I}_n = \mathrm{I}_r \quad 对于\ n \geqslant r.$$

设 I_n 的并集为 I, 这是一个理想. 由于 K 是主理想整环, 对于某个 $x \in \mathrm{K}$ 我们有 $\mathrm{I} = (x)$. 由于 x 属于 I_n 的并集, 至少对于一个指标 r 有 $x \in \mathrm{I}_r$. 但是那么显然 I_r 含有 x 的倍元, 故有 $\mathrm{I}_r \supset \mathrm{I}$, 并且对于 $n \geqslant r$ 我们得到

$$\mathrm{I}_r \subset \mathrm{I}_n \subset \mathrm{I} \subset \mathrm{I}_r,$$

由此推知 $\mathrm{I}_n = \mathrm{I}_r$, 这就证明了引理 3.

5. 素元的性质

下列结果包含素元的几个有用的特征 (其中包括素元的定义, 以便得到尽可能完整的表述):

定理 6　*设 p 是主理想整环 K 的一个非零且非可逆的元素, 则以下性质是等价的:*

a) *p 是素元;*

b) *p 的所有因子或是可逆的, 或有形式 pu, 其中 u 是可逆的;*

c) *理想 $\mathrm{I} = (p)$ 是极大的 (即 $\mathrm{I} \neq \mathrm{K}$, 并且 K 的仅有的包含 I 的理想是 I 和 K);*

d) *如果 p 整除 K 的元素的一个乘积, 则 p 整除这个乘积的至少一个因子.*

断言 a) 和 b) 的等价性正是素元的定义. 为了指出 b) 和 c) 的等价性, 将考虑一个包含 I 的理想 J. 由于 K 是主理想整环, 我们有对于某个 $x \in \mathrm{K}$, $\mathrm{J} = (x)$, 而说 J 包含 I 意味着 x 整除 p; 进而, 关系 $\mathrm{J} = \mathrm{K}$ 意味着 x 是可逆的, 而关系 $\mathrm{J} = \mathrm{I}$ 意味着 $x = pu$, 这里 u 是可逆的. b) 和 c) 的等价性立刻由这些注释得到.

最后指出性质 d) 也刻画了素元的特征. 假定 p 是素元, 并且 $x_1, \cdots, x_n$ 是 K 的非零元素, 使得 p 整除乘积 $x_1 \cdots x_n$. 为了指出 p 至少整除一个 x_i, 写出

$$x_1 \cdots x_n = (x_1 \cdots x_{n-1})x_n,$$

而归纳推理显然指出只需考察 $n = 2$ 的情形, 即证明如果 p 整除 xy, 则 p 或整除 x, 或整除 y. 根据定理 2, 所有事情归结为证明对于任意 $x \neq 0 \in \mathrm{K}$, 或者 p 整除 x, 或者 p 和 x 互素. 考虑由 p 和 x 生成的理想 (p, x), 它显然包含理想 (p), 根据已经确立的断言 c), 仅有两种可能情形: 或者

$$(p, x) = \mathrm{K},$$

那么 p 和 x 是互素的, 或者

$$(p, x) = (p),$$

在这种情形, x 属于 (p), 故是 p 的倍元.

我们已经证明了 K 的所有素元满足 d). 反之, 考虑一个非零并且非可逆的满足 d) 的元素, 要证明它是素元. 根据定理 5, 所考虑的元素 p 可以写成

$$p = p_1 \cdots p_n,$$

其中 p_i 是素元. 根据 d), p 至少整除一个 p_i, 由于 p_i 是素元, 而 p 是非可逆的, 故有 $p = up_i$, 这里 u 是可逆的, 因此 p 是素元, 这就结束了定理的证明.

推论 设 $x, y_1, \cdots, y_n$ 是主理想整环 K 的非零元素, 并且假定 x 与每个 y_i 互素. 则 x 与乘积 $y_1 \cdots y_n$ 互素.

事实上, 假定存在 x 和 $y_1 \cdots y_n$ 的一个非可逆的公因子 d. 由于 d 是不可逆的, 故它是素元的乘积, 从而是至少一个素元的倍元, 因此存在一个素元 p 同时整除 x 和 $y_1 \cdots y_n$. 根据定理 6, d), 存在一个 i, 使得 p 整除 p_i, 这与 x 和 y_i 的互素矛盾, 并且证明了推论.

不言而喻假定 $\mathrm{K} = \mathbf{Z}$ 丝毫不能简化前面的证明.

6. 素因子分解的唯一性

称两个素元 p' 和 p'' **是相伴的**, 如果存在 K 的一个可逆元 u, 使得 $p'' = up'$. 这显然意味着理想 (p') 和 (p'') 是相等的. 当 $\mathrm{K} = \mathbf{Z}$ 时这就是说 $p'' = \pm p'$.

定理 7 设 x 是主理想整环 K 的一个非零且非可逆的元素. 又设

$$x = p_1' \cdots p_m' = p_1'' \cdots p_n''$$

是 x 的 K 的素元乘积的两个分解. 则有 $m = n$, 并且存在指标 $1, \cdots, n$ 的一个置换, 使得 p_i' 与 $p_{\sigma(i)}''$ 是相伴的.

换句话说, 给定一个分解后, 对于这个分解进行如下操作可以得到所有其他的分解: (1) 改变素因子的次序, (2) 每个素因子乘以 K 的一个可逆元 (这些可逆元的选取要使得它们的乘积等于 1, 以便不改变所考虑的素元的乘积). 当 $\mathrm{K}=\mathbf{Z}$ 时, 经常只考虑正的素数, 这就取消了第二种可能性, 因为这时与 p 相伴的仅有的素数是 $-p$.

为了证明定理 7, 考虑关系

$$p_1'\cdots p_m' = p_1''\cdots p_n''. \tag{5}$$

p_1' 整除 $p_1''\cdots p_n''$, 并且是素元, 因此整除每个 p_j''. 改变第二个乘积中因子的次序, 可以假定 p_1' 不整除 p_1''. 但由于 p_1'' 是素元, 故存在可逆的 u_1 使得

$$p_1'' = up_1';$$

约去关系 (5) 中的 p_1' 即得

$$p_2'\cdots p_m' = up_2''\cdots p_n'';$$

重复如上推理, 推知 p_2' 相伴于 $p_j''(1\leqslant j\leqslant n)$ 之一. 如此推演下去, 显然就得到定理 7 (读者如果愿意正确地完成证明, 应当采用关于 m 的归纳推理).

在实际中, 经常以下列方式书写素因子分解. 一次性选定 K 的具有下列性质的素元的集合 P: 对于 K 的所有素元 p, 存在一个并且仅一个相伴于 p 的 $p'\in \mathrm{P}$ (为了构造一个这样的集合 P, 只需根据定理 6 的 c) 考虑 K 的极大理想的集合, 并且一次性地从这些理想中选定一个生成元. 如果 $\mathrm{K}=\mathbf{Z}$, 则对于每个极大理想 I 选择生成 I 的那个正素数). 设

$$x = p_1'\cdots p_n'$$

是非可逆的 $x\in \mathrm{K}$ 的一个素元乘积分解. 对于每个 i 可以写出 $p_i' = u_ip_i$, 这里 $p_i\in\mathrm{P}$ 并且 u_i 是可逆的, 由此显然得到

$$x = up_1\cdots p_n, \quad \text{其中 } u \text{ 是可逆的, 并且 } p_1,\cdots,p_n\in \mathrm{P}. \tag{6}$$

分解 (6) 不计因子的次序是唯一的 (因为 P 的元素不相等则不能是相伴的).

自然会碰到在分解 (6) 里 P 的一个元素重复出现的情形, 通过合并相同因子可以发现又一种形式的分解

$$x = up_1^{n_1}\cdots p_r^{n_r}, \tag{7}$$

其中的 $p_1,\cdots,p_r$ 是 P 的两两不等的元素, 而 n_i 是正的指数. 为了把这个结果书写成更统一的形式, 对于每个 $p\in\mathrm{P}$, 考虑分解 (6) 中等于 p 的因子的个数 (可能是零), 并且把它记作

$$v_p(x),$$

那么显然有

$$v_p(x) = 0 \quad \text{对于几乎所有的 } p\in \mathrm{P}. \tag{8}$$

换句话说, 仅对于有限个 $p \in \mathrm{P}$ 的元素有 $v_p(x) \geqslant 1$. 做了这些说明, 合并相同因子的集合所得到的分解 (7) 还可以写成形式

$$x = u \prod_{p \in \mathrm{P}} p^{v_p(x)}. \tag{9}$$

(出现在右端的乘积表面看有无穷多个因子, 但是根据 (8) 我们有

$$p^{v_p(x)} = 1 \quad \text{对于几乎所有的 } p \in \mathrm{P},$$

以致所提到的乘积仅含有有限个异于 1 的因子.)

7. 借助素因子分解求最大公因子和最小公倍

分解 (9) 允许十分简单地表述在环 K 内的整除性. 这一切建立在下列结果的基础之上:

引理 4 设 x 和 y 是环 K 的两个非零元. x 整除 y, 必须并且只需对于所有 $p \in \mathrm{P}$ 有

$$v_p(x) \leqslant v_p(y).$$

条件是充分的, 因为当它满足时, 我们有

$$x = u' \prod_{p \in \mathrm{P}} p^{v_p(x)},$$
$$y = u'' \prod_{p \in \mathrm{P}} p^{v_p(x)+n_p},$$

其中的整数

$$n_p = v_p(y) - v_p(x)$$

都是非负的并且几乎全部为零, 故有 $y = xz$, 这里

$$z = u'^{-1} u'' \prod p^{n_p}.$$

反之, 假定 $y = xz$, 利用 x 和 z 的素因子分解即得

$$y = u \prod_{p \in \mathrm{P}} p^{v_p(x)+v_p(z)},$$

其中的 $u \in \mathrm{K}$ 是可逆的. 由于 y 的分解成 P 的元素和可逆元的乘积的形式是唯一的, 我们发现 $v_p(y) = v_p(x) + v_p(z)$, 故对于所有的 $p \in \mathrm{P}$ 有 $v_p(y) \geqslant v_p(x)$, 这就结束了引理 4 的证明.

给定 K 的两个非零元素, 借助引理 4 构造它们的最大公因子和最小公倍是容易的. 事实上, 设 d 是 x 和 y 的最大公因子, 我们应该表述 x 和 y 的因子是 d 的因子, 而 x 和 y 的公因子根据引理 4 是满足

$$v_p(z) \leqslant v_p(x) \quad \text{和} \quad v_p(z) \leqslant v_p(y)$$

的 $z \in \mathrm{K}$, 即对于所有 $p \in \mathrm{P}$ 有

$$v_p(z) \leqslant \min(v_p(x), v_p(y)). \tag{10}$$

而 d 的因子是对于所有的 $p \in \mathrm{P}$ 满足

$$v_p(z) \leqslant v_p(d) \tag{11}$$

的 $z \in \mathrm{K}$. d 是 x 和 y 的最大公因子, 必须并且只需 (10) 和 (11) 是等价的, 即

$$v_p(z) = \min(v_p(x), v_p(y)) \quad \text{对于所有 } p \in \mathrm{P}, \tag{12}$$

我们这样就重新发现了经典的法则: d 的素元分解里 p 的指数等于在 x 和 y 的分解里 p 的指数的较小者.

类似的推理表明 x 和 y 的最小公倍 m 由下式给定:

$$v_p(m) = \max(v_p(x), v_p(y)) \quad \text{对于所有 } p \in \mathrm{P}. \tag{13}$$

即在环 **Z** 的情形的熟知法则推广到了所有主理想整环, 并且借助经典情形的同样的推理得以建立.



注 3 不言而喻, 形如

$$\prod_{p \in \mathrm{P}} p^{n_p}$$

的乘积仅当指数 n_p 是非负整数并且几乎全部为零时才有意义. 如果想利用关系 (12) 和 (13) 确定最大公因子和最小公倍, 就必须验证这些关系的右端满足所提出的条件. 如下确认这些条件. 设 X (对应的, Y) 是使得 $v_p(x) \neq 0$ (对应的, $v_p(y) \neq 0$) 的 $p \in \mathrm{P}$ 的集合, 即整除 x (对应的, y) 的 $p \in \mathrm{P}$ 的集合, X 和 Y 是有限集合, 故 $\mathrm{Z} = X \cup \mathrm{Y}$ 也是有限集合. 对于 $p \notin \mathrm{Z}$ 有 $v_p(x) = v_p(y) = 0$, 故

$$\max(v_p(x), v_p(y)) = \min(v_p(x), v_p(y)) = 0,$$

这就是要达到的结果.

8. 主理想整环上的分式的部分分式分解

设 K 是一个主理想整环, 而 F 是 §29 中定义的 K 的分式域. 设

$$x = a/b, \quad a, b \in \mathrm{K}, \quad b \neq 0$$

是 F 的一个非零元素. 保持前一个小节的记号, 可以写出

$$a = u \prod_{p \in \mathrm{P}} p^{v_p(a)}, \quad b = v \prod_{p \in \mathrm{P}} p^{v_p(b)},$$

由此得到

$$x = w \prod_{p \in \mathrm{P}} p^{v_p(x)}, \tag{14}$$

其中的 w 是 K 的一个可逆元, 而整数

$$v_p(x) = v_p(a) - v_p(b)$$

仍然是几乎全部为零, 不过可以是任意符号的. 容易验证 —— 读者作为习题可以证明 —— x 的分解 (14) 是唯一的.

现在对于 F 的元素建立完全另一种性质的分解, 下面的结果在分析中是有用的 (实或复系数的有理函数的原函数的积分法), 但是几乎完全没有其他用途.

定理 8 设

$$x = \frac{a}{p_1^{r_1} \cdots p_n^{r_n}}$$

是 F 的一个元素, 其中 $a \in \mathrm{K}$, r_i 是正整数, $p_i \in \mathrm{K}$ 是两两不相伴的素元. 则存在 K 的元素 $a_1, \cdots, a_n$, 使得

$$x = \frac{a_1}{p_1^{r_1}} + \cdots + \frac{a_n}{p_n^{r_n}}.$$

即 F 的所有元素是形如 a/p^n 的分式的和, 其中 $a \in \mathrm{K}$, $p \in \mathrm{K}$ 是素元, 并且 $n \geqslant 0$.

例 2 取 $\mathrm{K} = \mathbf{Z}$ 和

$$x = \frac{5}{18} = \frac{5}{2 \cdot 3^2};$$

则有

$$x = \frac{1}{2} - \frac{2}{9},$$

这是在这种情形的定理 8 所断言的结果.

定理显然是以下两个引理的推论:

引理 5 *假定*

$$x = a/b_1 \cdots b_n,$$

其中 K 的元素 b_i 是两两互素的. 则存在 $a_i \in \mathrm{K}$. 使得

$$x = a_1/b_1 + \cdots + a_n/b_n.$$

引理 6 设 $p_1, \cdots, p_n$ 是 K 的两两不相伴的素元. 则对于所有正整数 $r_1, \cdots, r_n$, 元素

$$b_1 = p_1^{r_1}, \cdots, b_n = p_n^{r_n}$$

是两两互素的.

先证明引理 6. 设 d 是 b_i 和 b_j $(i \neq j)$ 的一个公因子, 如果 d 不是可逆的, 则它有一个素因子 p, p 整除 b_i 和 b_j. 但是如果 p 整除

$$p_i^{r_i} = p_i \cdots p_i,$$

它将整除 p_i, 因而和 p_i 相伴. 于是如果 b_i 和 b_j 不是互素的, 将存在 K 的一个素元 p 同时与 p_i 和 p_j 相伴, 这跟对于 $i \neq j$, p_i 和 p_j 不相伴的假定矛盾, 由此得到引理 6.

现在证明引理 5, 首先对于 $n = 2$ 证明. 由于 p_1 和 p_2 是互素的, 存在 $u_1, u_2 \in \mathrm{K}$, 使得

$$u_1 b_1 + u_2 b_2 = 1,$$

于是有

$$\frac{a}{b_1 b_2} = \frac{a(u_1 b_1 + u_2 b_2)}{b_1 b_2} = \frac{a u_2}{b_1} + \frac{a u_1}{b_2},$$

对于这个情形这就证明了引理 5. 为了证明一般情形, 要关于 n 进行归纳推理. 由于(根据定理 6 的推论) b_n 与 $b_1 \cdots b_{n-1}$ 互素, 又由于对于两个因子的乘积已经证明了引理, 可以写出

$$x = \frac{a'}{b_1 \cdots b_{n-1}} + \frac{a_n}{b_n},$$

再对于右端的第一个分式利用归纳假设就得到对于 x 所找的分解.

§31 习题

1. 设 $x_1, \cdots, x_n$ 是一个主理想整环 K 的非零元素, 而 d 是它们的最大公因子. 选择 $u_1, \cdots, u_n \in \mathrm{K}$, 使得 $u_1 x_1 + \cdots + u_n x_n = d$. 证明 $u_1, \cdots, u_n$ 是互素的.

¶2. 设 M 是一个主理想整环 K 上的有限生成的自由模, 而 a 是 M 的一个非零元素. 证明下列关于 a 的五个性质是等价的:

(i) a 是 M 的一个基的一部分;

(ii) 存在 M 上的一个线性型 f, 使得 $f(a) = 1$;

(iii) a 关于 M 的一个基的坐标是互素的; **a 关于 M 的所有基的坐标是互素的;**

(iv) 如果对于一个 $u \in \mathrm{K}$ 和 M 内的一个 $x \neq 0$ 有 $a = ux$, 那么 u 是可逆的;

(v) 如果 $ux = va$, 其中 $u, v \in \mathrm{K}$, 并且 $x \in \mathrm{M}$ 非零, 那么 v 是 u 的倍元. (利用 §18 的习题 2 和 §29 的习题 11, h).)

满足这些条件的向量 $a \in \mathrm{M}$ 称为**本原的**.

3. 设 M 是主理想整环 K 上的有限生成的自由模. 证明所有 $x \in \mathrm{M}$ 是至少 M 的一个本原向量的倍元. 取 $\mathrm{K} = \mathbf{Z}, \mathrm{M} = \mathbf{Z}^4$ 和 $x = (126, 210, 168, 504)$, 求这样的一个本原向量.

4. 设 $a_1, \cdots, a_n$ 是主理想整环 K 的元素. 存在一个其第一行 (对应的, 列) 正好是 $a_1, \cdots, a_n$ 的矩阵

$$\mathrm{U} \in \mathrm{GL}(n, \mathrm{K}),$$

必须并且只需 $a_1, \cdots, a_n$ 是互素的. 这时就可以选取 U, 使得 $\det(\mathrm{U}) = 1$, 即可以假定

$$\mathrm{U} \in \mathrm{SL}(n, \mathrm{K}).$$

5. 构造一个其第一列是 2, 3, 4 的矩阵 $\mathrm{U} \in \mathrm{SL}(3, \mathbf{Z})$.

6. 构造一个其第二列是 2, 3, 4 的矩阵 $\mathrm{U} \in \mathrm{SL}(3, \mathbf{Z})$.

¶7. 设 M 是主理想整环 K 上的一个有限生成的自由模.

a) 证明, 如果 a 是 M 的一个非零元, 那么 a 关于 M 的一个基的坐标的最大公因子不依赖这个基的选取. 这个结果的 "几何" 解释是什么 (参见习题 3) ?

b) 设 M' 是 M 的一个非零子模. 选择 M 一个基, 并且考虑由 M' 的所有元素的所有坐标生成的 K 的理想. 证明这个理想不依赖基的选取. 证明它是由 M' 的任意生成集的元素的坐标生成的. [这个理想或它的任何一个生成元称为 M' 在 M 内的**第一不变因子**, 见下一个习题.]

¶¶8. 设 M 是主理想整环 K 上的一个有限生成的自由模, 而 M' 是 M 的一个非零子模. 把 M 和 M' 的秩 (一个基的元素的个数) 记为 n 和 r. 我们打算证明以下结果: *存在 M 的一个基 $a_1, \cdots, a_n$ 和 K 的元素 $d_1, \cdots, d_r$, 使得向量 $d_1 a_1, \cdots, d_r a_r$ 组成 M' 的一个基, 并且对于 $1 \leqslant i \leqslant r-1$ 有 d_i 整除 d_{i+1}.*

a) 证明, 对于 M 上的所有线性型, 集合 $f(\mathrm{M}') \subset \mathrm{K}$ 是 K 的一个理想.

b) 证明存在 M 上的一个线性型 f_1, 使得对于 M 上的所有线性型 f 有关系

$$f_1(\mathrm{M}') \subset f(\mathrm{M}') \quad \text{蕴含} \quad f_1(\mathrm{M}') = f(\mathrm{M}').$$

那么就有

$$f_1(\mathrm{M}) = \mathrm{K}.$$

c) 取满足 b) 的 f_1, 令

$$f_1(\mathrm{M}') = (d_1),$$

并且选取一个向量 $u_1 \in \mathrm{M}'$, 使得

$$f_1(u_1) = d_1.$$

证明对于 M 上的所有线性型 f

$$f(u_1) \in (d_1)$$

(令 $f(u_1) = d$, 证明存在 f 和 f_1 的一个线性组合 g, 使得 $g(u_1)$ 是 d 和 d_1 的一个最大公因子).

d) 由 c) 推出对于使得 $f_1(e_1) = 1$ 的一个向量 $e_1 \in \mathrm{M}$ 有

$$u_1 = d_1 e_1.$$

e) 证明 M 是 e_1 生成的子模和 $\mathrm{Ker}(f_1)$ 的直和; 而 M' 是 u_1 生成的子模和 $\mathrm{M}'\cap\mathrm{Ker}(f_1)$ 的直和. 证明对于所有 f 有 $f(\mathrm{M}')\subset f_1(\mathrm{M}')$.

f) 用关于 n 的归纳法完成证明.

¶¶9. 设 A 是其元素在一个主理想整环 K 内的 n 行 p 列的矩阵. 借助习题 8 证明存在矩阵

$$\mathrm{U}\in\mathrm{GL}(n,\mathrm{K}) \quad 和 \quad \mathrm{V}\in\mathrm{GL}(p,\mathrm{K})$$

使得

$$\mathrm{UAV}=\begin{pmatrix} d_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix},$$

其中的 $d_1,\cdots,d_r$ 是 K 的非零元素, 每一个整除下一个. $d_1,\cdots,d_r$ 称为矩阵 A 的**不变因子**. 将看到 (习题 11) 理想 $(d_1),\cdots,(d_r)$ 完全由 A 确定.

¶¶10. 证明所有有限生成交换群同构于一个群 $\mathbf{Z}^q$ 和有限个循环群 $\mathbf{Z}/d_1\mathbf{Z},\cdots,\mathbf{Z}/d_r\mathbf{Z}$ 的直积, 其中 d_i 整除 $d_{i+1}(1\leqslant i\leqslant r-1)$. (注意到一个有限生成 K-模, 其中 K 是任意一个环, 同构于商 M/M', 其中 M 是有限生成的自由的, 并且应用习题 8.) 怎样把这个结果推广到任意一个主理想整环?

¶¶11. 继续采取习题 8 的假设和记号并且利用它的结果.

a) 设 $j_1,\cdots,j_h$ 是整数, 满足条件

$$1\leqslant j_1<\cdots<j_h\leqslant r,$$

证明 $d_1\cdots d_h$ 整除 $d_{j_1}\cdots d_{j_h}$.

b) 设 h 使得 $1\leqslant h\leqslant r$, 证明如果 f 是一个交错 h 重线性型, 则对于任意 $x_1,\cdots,x_h\in\mathrm{M}'$, 乘积 $d_1\cdots d_h$ 整除 $f(x_1,\cdots,x_h)$. 证明还可以选择 f 和 $x_1,\cdots,x_h\in\mathrm{M}'$, 使得

$$d_1\cdots d_h=f(x_1,\cdots,x_h).$$

由此推出 $d_1\cdots d_h$ 是形式为 $f(x_1,\cdots,x_h)$ 的 K 的元素的最大公因子, 并且由此推断出理想 (d_i) 完全由模和子模 M' 确定 (即不依赖习题 8 中所构造的基的选择).

c) 设 $(a_i)_{1\leqslant i\leqslant n}$ 是 M 的任意一个基, 而 $(b_j)_{1\leqslant j\leqslant p}$ 是 M' 的任意一个生成元组 $(b_j)_{1\leqslant j\leqslant p}$. 用 A 表示由 b_j 关于 M 的基 (a_i) 的坐标组成的 (n 行 p 列) 矩阵.

证明对于 $1\leqslant h\leqslant r$, 元素 $d_1\cdots d_h$ 是矩阵 A 的 h 阶子式的最大公因子.

d) 由此推出习题 9 的因子 $d_1,\cdots,d_r$ 用同样的方法计算.

e) 设 A 和 B 是其元素在一个主理想整环 K 内的两个 n 行 p 列的矩阵. A 和 B 是**等价的** (即存在可逆矩阵使得 $\mathrm{B}=\mathrm{UAV}$) 必须并且只需 A 和 B 有同样的秩和同样的不变因子. (注意代替 A 的不变因子, 经常引进**初等因子**

$$e_1=d_1, e_2=d_2/d_1,\cdots,e_r=d_r/d_{r-1},$$

以便表达这个结果.)

¶¶12. 设 K 是一个主理想整环, 而

$$a_j = (\alpha_{1j}, \cdots, \alpha_{nj}) \quad (1 \leqslant j \leqslant p)$$

是 K^n 的 p 个元素. 它们组成 K^n 的一个基的一部分, 必须并且只需由给定向量的分量组成的矩阵

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1p} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{np} \end{pmatrix}$$

的 p 阶子式是*互素的* (特别, 它们不全为零).

¶¶13. 设 K 是一个主理想整环, 而 n 和 p 是两个整数, $1 \leqslant p \leqslant n$. 设 A 是一个系数在 K 内的 n 行 p 列的矩阵. 为了能够把 A 补充为一个在 K 上*可逆的* n 阶方阵, 必须并且只需 A 的所有 p 阶子式的最大公因子是 1.

14. 求形式为

$$\begin{pmatrix} 1 & 4 & * \\ 2 & 5 & * \\ 1 & 6 & * \end{pmatrix}$$

并且其行列式为 1 的其元素为有理整数的所有矩阵.

¶15. 设 A 是元素在一个交换环 K 内的矩阵. 对于 A 的一个变换, 它或者是交换 A 的两行 (对应的, 列), 或者是一行 (对应的, 一列) 加上其余各行 (对应的, 列) 的一个线性组合, 或者是一行 (对应的, 一列) 乘以 K 的一个可逆元, 称为对于 A 的**初等变换**.

a) 证明从 A 经过一系列初等变换得到的矩阵与 A 是等价的 (即其形式为 UAV, 而 U,V 在 K 上是可逆的).

b) 假定 $\mathrm{K} = \mathbf{Z}$, 并且 $\mathrm{A} \neq 0$. 设 d_1 是具有下列性质的最小正整数: 存在一个从 A 经过一系列初等变换得到的矩阵, 使得 d_1 是它的一个元素. 证明那么就存在一个从 A 经过一系列初等变换得到的矩阵, 其形式为

$$\begin{pmatrix} d_1 & 0 \\ 0 & \mathrm{A}_1 \end{pmatrix}$$

(A_1 的行数和列数比 A 少 1), 并且 A_1 的元素都是 d_1 的倍数.

c) 由此得到 (在 $\mathrm{K} = \mathbf{Z}$ 时) 习题 9 (同样还有习题 8) 的一个新的证明和把元素在 $\mathbf{Z}$ 内的矩阵化成习题 9 的典范形式的一个实际方法.

d) 应用这个方法到下列矩阵:

$$\begin{pmatrix} 0 & 2 & 4 & -1 \\ 6 & 12 & 14 & 5 \\ 0 & 4 & 14 & -1 \\ 10 & 6 & -4 & 11 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 6 & -9 & -3 \\ 12 & 24 & 9 & 9 \\ 30 & 42 & 45 & 27 \\ 66 & 78 & 81 & 63 \end{pmatrix},$$

$$\begin{pmatrix} 17 & -28 & 45 & 11 & 39 \\ 24 & -37 & 61 & 13 & 50 \\ 25 & -7 & 32 & -18 & -11 \\ 31 & 12 & 19 & -43 & -55 \\ 42 & 13 & 29 & -55 & -68 \end{pmatrix}.$$

16. 设 A 是元素在一个任意的交换环 K 内的矩阵, 而 B 是等价于 A (即其形式为 UAV, U 和 V 在 K 上是可逆的) 的一个矩阵. 证明, 对于不超过 A 的行数和列数的所有整数 p, 由 A 的 p 阶子式生成的 K 的理想等于由 B 的 p 阶子式生成的 K 的理想.

¶17. 设 A 是元素在一个主理想整环 K 内的 n 阶方阵. 证明存在一个矩阵 $\mathrm{U} \in \mathrm{GL}(n, \mathrm{K})$, 使得 UA 是三角矩阵 (利用习题 1 和 4, 并且关于 n 进行归纳推理). 几何解释如何?

¶¶18. 考虑系数、右端和未知元在主理想整环 K 内的线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1p}x_p = b_1, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1 + \cdots + a_{np}x_p = b_n. \end{cases}$$

证明这个方程组至少具有一个解, 必须并且只需两个矩阵

$$\mathrm{A} = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}, \qquad \mathrm{B} = \begin{pmatrix} a_{11} & \cdots & a_{1p} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{np} & b_n \end{pmatrix}$$

有同样的秩 r, 此外 A 的 r 阶子式的最大公因子等于 B 的 r 阶子式的最大公因子.

¶¶19. 设 K 是一个主理想整环, f 是模 K^n 上的 p 重线性型, 而 d 是 f 关于典范基的坐标的一个最大公因子. 证明存在 $x_1, \cdots, x_p \in \mathrm{K}^n$, 使得

$$f(x_1, \cdots, x_p) = d.$$

20. 证明所有大于 0 小于 1 的有理数可以用唯一的一种方式写成有限个形式

$$a/p^n,$$

的分数的和, 其中的 p 为素数, 并且 $1 \leqslant a \leqslant p-1$.

对于数

$$\frac{1887}{5400}, \quad \frac{122}{1323}$$

进行这种分解.

¶¶21. 设 K 是一个交换整环. 说 $p \in \mathrm{K}$ 是**不可约的**, 如果它不是可逆的, 并且它没有显然的因子 (即 K 的可逆元和 up, 其中 $u \in \mathrm{K}$ 是可逆的) 以外的因子. 称 K 是**唯一因子分解整环**, 如果它有下列两个性质:

(UFD1) K 的所有非零非可逆的元素是有限个不可约元素的乘积.

(UFD2) 如果 $p_1 \cdots p_r = q_1 \cdots q_s$ 是 K 的不可约元素的乘积, 则有 $r = s$, 并且可以改变 q_j 的次序, 使得有

$$\mathrm{K}_{p_i} = \mathrm{K}q_i, \quad 对于\ 1 \leqslant i \leqslant r.$$

(换句话说, $q_i = u_i p_i$, u_i 是可逆的).

这些性质表明 K 的所有元素可以用本质上同一种方式写成 K 的不可约元素的乘积.

a) 证明所有主理想整环是唯一因子分解整环 (其逆是错误的; 参见 §32, 习题 31).

b) 证明在唯一因子分解整环的定义中, 可以把条件 (UFD2) 换成

(UFD3) 如果 K 的不可约元素 p 整除一个乘积 xy, 则它至少整除这个乘积的两个因子中的一个.

c) 证明还可以把 (UFD2) 换成

(UFD4) 对于 K 的所有不可约元素 p, 理想 $\mathrm{K}p$ 是素理想.

d) 设 K 是一个唯一因子分解整环. 证明对于任意 $x, y \in \mathrm{K}$, 存在一个 $d \in \mathrm{K}$, 使得 x 和 y 的公因子是 d 的因子, 反之亦然 (称 d 是 x 和 y 的**最大公因子**), 并且如果忽略 K 的一个可逆元因子的差别, d 是唯一的.

e) 设 K 是一个唯一因子分解整环, L 是它的分式域, 而 $\mathfrak{p}$ 是 K 的一个不可约元 p 生成的素理想. 证明局部环 $\mathrm{K}_{\mathfrak{p}}$ [§8, 习题 7, g)] 是 L 的离散赋值环 (§8, 习题 6).

f) 如果一个交换整环既是唯一因子分解整环, 又是 Dedekind 的, 则它是主理想整环.

g) 设 K 是交换的 Noether 整环. 证明 K 的所有非可逆元是不可约元的乘积, 换句话说, K 满足 (UFD1). [但是 Noether 环未必是唯一因子分解整环, 即不可约元因式分解可能不是唯一的: 比如, 取一个不是主理想整环的 Dedekind 整环; 而这一个现象尤其来自代数数域的整数环, 并且长期以来阻碍了这些环的研究的进展. 直到 Dedekind, 他第一个认识到在这种情形重要的概念是素理想的概念而非不可约元素的概念, 而这与有理整数的迷惑人的类似所显示的情形相反.]

§32 多项式除法

1. 一个未定元的多项式除法

设

$$f(\mathrm{X}) = a_0 + a_1\mathrm{X} + \cdots + a_n\mathrm{X}^n$$

是系数在交换环 K 内的一个未定元的多项式, 称 f 的最高次数项的系数为 f 的**首项系数** (假定 $f \neq 0$). 如果 f 是 n 次的, f 的首项系数就是 X^n 的系数. 如果一个多项式 f 的首项系数是 K 的可逆元, 则称 f 是**归一多项式**. 如果 K 是一个域, 则所有多项式 $f \neq 0$ 都是归一多项式.

以下结果是多项式类似于有理整数带余除法的相应结果:

定理 1 *设 K 是交换环, 而 g 是系数在 K 内的一个未定元的归一多项式. 则对于所有多项式 $f \in \mathrm{K}[\mathrm{X}]$, 存在多项式 $q, r \in \mathrm{K}[\mathrm{X}]$, 满足关系*

$$f = gq + r, \quad d^\circ(r) < d^\circ(g); \tag{1}$$

并且多项式 q 和 r 是唯一的.

设

$$g(\mathrm{X}) = b_0 + b_1\mathrm{X} + \cdots + b_n\mathrm{X}^n,$$

其中的 b_n 是可逆的. 上式两端乘以 b_n^{-1}, 显然就可以归结为 g 取以下形式的情形:

$$g(\mathrm{X}) = \mathrm{X}^n + b_{n-1}\mathrm{X}^{n-1} + \cdots + b_0. \tag{2}$$

以下我们就假定 g 具有形式 (2).

考虑形式为 $f - gq$ (其中 f 和 g 给定, 而 q 变动) 的所有多项式, 它们的次数为非负整数或 $-\infty$ (如果 f 是 g 的倍式, 则在这种情形定理必然是平凡的). 因此可以选择 q, 使得 $f - gq$ 的次数最小, 于是对于所有 $q' \in \mathrm{K}[\mathrm{X}]$ 有

$$d^\circ(f - gq) \leqslant d^\circ(f - gq'). \tag{3}$$

令

$$f = gq + r, \tag{4}$$

事情归结为确认

$$d^\circ(r) < d^\circ(g). \tag{5}$$

假定不然, 令

$$r(\mathrm{X}) = c_{n+k}\mathrm{X}^{n+k} + c_{n+k-1}\mathrm{X}^{n+k-1} + \cdots,$$

由于假定关系 (5) 不成立, 故其中的 $c_{n+k} \neq 0$, 并且 $k \geqslant 1$. 显然 r 和多项式 $c_{n+k}\mathrm{X}^k g(\mathrm{X})$ 有相同的首项系数 c_{n+k}, 因此可以写出

$$r(\mathrm{X}) = c_{n+k}\mathrm{X}^k g(\mathrm{X}) + r'(\mathrm{X}), \quad 其中\ d^\circ(r') < d^\circ(r). \tag{6}$$

此时 (4) 改写为

$$f(\mathrm{X}) = [q(\mathrm{X}) + c_{n+k}\mathrm{X}^k]g(\mathrm{X}) + r'(\mathrm{X}),$$

令

$$q'(\mathrm{X}) = q(\mathrm{X}) + c_{n+k}\mathrm{X}^k,$$

则有

$$f = qq' + r', \quad 其中\ d^\circ(r') < d^\circ(r).$$

这最后的不等式写成 $d^\circ(f - gq') < d^\circ(f - gq)$, 这与 (3) 矛盾, 并且证明了 (5).

我们已经证明了至少存在一对多项式 q, r 满足 (1), 还需要指出至多存在一对. 我们考虑形式如下的两个关系

$$f = gq_1 + r_1 = gq_2 + r_2,$$

其中有

$$d^\circ(r_1) < d^\circ(g), \quad d^\circ(r_2) < d^\circ(g). \tag{7}$$

由上得到

$$g(q_1 - q_2) = r_2 - r_1, \tag{8}$$

一切归结为指出 $q_1 = q_2$. 而根据 (7), 我们有 $d^\circ(r_2-r_1) < d^\circ(g)$, 于是关系 $q_2-q_1=0$ 将是以下结果的推论:

引理 1 如果 g 是一个归一多项式, 而 q 是一个非零多项式, 则有

$$d^\circ(gq) = d^\circ(g) + d^\circ(q) \geqslant d^\circ(g).$$

令

$$g(\mathrm{X}) = b_n\mathrm{X}^n + \cdots + b_0, \quad q(\mathrm{X}) = c_m\mathrm{X}^m + \cdots + c_0,$$

其中的 b_n 是可逆的, 而 c_m 不是零. 我们发现在 gq 中 X^{m+n} 的系数等于 b_nc_m. 由于 b_n 是可逆的, 这个乘积只有 $c_m = 0$ 才会等于零, 这是不可能的. 由此得到引理.

至此定理 1 证明完毕.

在定理 1 的表述中 (如果 K 是一个域, 定理 1 仅假设 $g \neq 0$ 即可), 称 q 是 f **除以 g 的商式**, 而 r 称为 f **除以 g 的余式**. 定理 1 的证明引导出一个计算它们的实际方法. 事实上, 令

$$f(\mathrm{X}) = a_m\mathrm{X}^m + \cdots + a_0, \quad g(\mathrm{X}) = b_n\mathrm{X}^n + \cdots + b_0,$$

其中 a_m 非零, 而 b_n 是可逆的. 如果 $m < n$, 除法平凡地实施: 只需取 $q = 0, r = f$. 如果 $m \geqslant n$, 显然有

$$f(\mathrm{X}) = a_mb_n^{-1}\mathrm{X}^{m-n}g(\mathrm{X}) + f_1(\mathrm{X}) = c_k\mathrm{X}^kg(\mathrm{X}) + f_1(\mathrm{X}), \tag{9}$$

其中

$$d^\circ(f_1) \leqslant d^\circ(f) - 1.$$

如果 f_1 的次数低于 g 的次数, 则除法结束, 并且在这种情形可以精确地写出

$$q(\mathrm{X}) = a_mb_n^{-1}\mathrm{X}^{m-n}, \quad r(\mathrm{X}) = f_1(\mathrm{X});$$

如果不然则有 $d^\circ(f_1) \geqslant d^\circ(g)$, 对于 f_1, 像 f 一样如法炮制, 可以写出

$$f_1(\mathrm{X}) = c_h\mathrm{X}^hg(\mathrm{X}) + f_2(\mathrm{X}), \tag{10}$$

其中

$$d^\circ(f_2) < d^\circ(f_1).$$

组合 (9) 和 (10) 即得

$$f(\mathrm{X}) = (c_k\mathrm{X}^k + c_h\mathrm{X}^h)g(\mathrm{X}) + f_2(\mathrm{X}), \quad d^\circ(f_2) < d^\circ(f) - 2;$$

如果 f_2 的次数低于 g 的次数, 则问题解决, 否则写出

$$f_2(\mathrm{X}) = c_l\mathrm{X}^l g(\mathrm{X}) + f_3(\mathrm{X}),$$

其中

$$d^\circ(f_3) < d^\circ(f_2),$$

于是有

$$f(\mathrm{X}) = (c_k\mathrm{X}^k + c_h\mathrm{X}^h + c_l\mathrm{X}^l)g(\mathrm{X}) + f_3(\mathrm{X}), \quad d^\circ(f_3) < d^\circ(f) - 3;$$

显然这个过程在有限步之后将会终止, 我们就会得到 f 除以 g 的商式和余式.

例 1 取 $f(\mathrm{X}) = \mathrm{X}^6 - \mathrm{X}^4 - \mathrm{X}^2 + 1, \quad g(\mathrm{X}) = \mathrm{X}^3 - 1$. 我们如下安排运算:

$$\begin{array}{rrrrr|l}
\mathrm{X}^6 - \mathrm{X}^4 & & -\mathrm{X}^2 & & +1 & \mathrm{X}^3 - 1 \\ \cline{6-6}
-\mathrm{X}^4 & +\mathrm{X}^3 & -\mathrm{X}^2 & & +1 & \mathrm{X}^3 - \mathrm{X} + 1 \\
& \mathrm{X}^3 & -\mathrm{X}^2 & -\mathrm{X} & +1 & \\
& & -\mathrm{X}^2 & -\mathrm{X} & +2 &
\end{array}$$

这里商式是 $\mathrm{X}^3 - \mathrm{X} + 1$, 而余式是 $-\mathrm{X}^2 - \mathrm{X} + 2$. 我们所采用的计算方法和对于整数的除法是相同的.

2. 一个未定元的多项式环中的理想

定理 1 的最重要的一个推论如下:

定理 2 *设 K 是一个交换域, 则系数在 K 内的一个未定元的多项式环 K[X] 是主理想环.*

环 K[X] 是交换的整环 (§27, 定理 1). 设 I 是 K[X] 的一个理想, 为了指出这是一个主理想, 可以假定它没有缩减为 0. 在 I 的非零元中, 选择一个次数最低的多项式 f, 那么显然有关系

$$r \in \mathrm{I} \quad \text{并且} \quad d^\circ(r) < d^\circ(f) \quad \text{蕴含} \quad r = 0. \tag{11}$$

现在设 g 是 I 的一个元素, 由于 K 是一个域并且 f 不是零, 定理 1 指出有等式

$$g = fq + r, \quad d^\circ(r) < d^\circ(f).$$

由于 I 含有 f 和 g, 它也含有 $g - fq$, 即含有 r, 于是利用 (11) 即得 $r = 0$. 如此看来, I 的所有元素 g 都是 f 的倍元, 而 f 的所有倍元显然在 I 内, 这就完成了证明.

读者应当把这里的证明与 §7 例 8 做比较, 两种情形所用的方法是类似的.

设 $\mathrm{I} = (f)$ 是环 K[X] 的一个理想, 多项式 f 如果不计它可以乘以 K[X] 的可逆元则是唯一的. 在这种情况下有下列结果:

引理 2 设 K 是一个交换整环, K[X] 的可逆元就是 K 的可逆元.

显然 K 的可逆元是 K[X] 的可逆元. 反之设

$$f(\mathrm{X}) = a_n\mathrm{X}^n + \cdots + a_0, \quad a_n \neq 0$$

是 K[X] 的一个可逆元, 设

$$g(\mathrm{X}) = b_m\mathrm{X}^m + \cdots + b_0, \quad b_m \neq 0$$

是它的逆元, 求出 f 乘以 g 的乘积,

$$1 = a_nb_m\mathrm{X}^{m+n} + \cdots,$$

没有写出的项的次数低于 $m+n$. 由于 K 是一个整环, 应当有 $a_nb_m \neq 0$, 故 $m+n=0$, 即 f 缩减为 K 的元素 a_n, 考虑到 $a_nb_m = 1$, a_n 是可逆的, 这就完成了证明.

还可以直接写出 (§27, 定理 1)

$$d^\circ(f) + d^\circ(g) = d^\circ(fg) = 0,$$

这就迫使 f 和 g 缩减为 K 的元素.

引理 2 表明, 如果 f 是 K[X] 的一个理想 I 的一个生成元 (重新假定 K 是一个域), 则 I 的一个其他生成元由 f 乘以非零常元 (即 K 的一个元素) 而得到. 在特殊情形, f 乘以它的首项系数的逆元, 就可以调整 f 的首项系数为 1, 那么 f 就是 I 的唯一的具有这个性质的生成元. 我们发现像在经典情形环 **Z** 一样, 对于所考虑的环 K[X] 的每个非零理想存在选择生成元的 "典范" 方法.

定理 2 对于解决相对初等的问题是非常重要和有用的, 但是一旦接触到 "代数几何" (参见后面的注 3) 就需要许多比定理 2 更一般的结果. 这种观点下的基本的结果是: 如果 K 是交换的 Noether 环 (即其所有理想是有限生成的), 那么环 K[X] 仍然是交换的 Noether 环. 在本节的习题 27 中将会发现这个结果的一个简单的证明. 关于 n 进行归纳推理, 就可以推导出更一般的结果: K 是 Noether 环, 那么环 $\mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$ 仍然是 Noether 环. 其一个特例是: 如果 K 是一个域, 那么对于任何 n, $\mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$ 是 Noether 环. 这些属于 Hilbert 的结果, 解释了为什么涉及系数在一个域内的多项式计算总可以执行 "有限步" 即结束, 但其重要性远远超过了多少这个有些玄奥的注释.

3. 几个多项式的最大公因式和最小公倍式

当 K 是一个域时, 定理 2 使我们能够把前一节的结果应用到环 K[X].

特别考虑系数在 K 内的一个未定元的非零多项式 $f_1, \cdots, f_n$, 那么这些多项式拥有一个**最大公因式** d. 我们有下列结果:

定理 3 设 $d, f_1, \cdots, f_n$ 是系数在 K 内的一个未定元的非零多项式, 则下列性质是等价的:

a) d 是 $f_1, \cdots, f_n$ 的最大公因式;

b) $f_1, \cdots, f_n$ 的公因式是 d 的因式;

c) 一个多项式可以写成形式

$$u_1(\mathrm{X})f_1(\mathrm{X}) + \cdots + u_n(\mathrm{X})f_n(\mathrm{X}), \quad u_i \in \mathrm{K}[\mathrm{X}] \quad (1 \leqslant i \leqslant n),$$

必须并且只需它是 d 的倍式;

d) $d(\mathrm{X})$ 可以写成形式

$$u_1(\mathrm{X})f_1(\mathrm{X}) + \cdots + u_n(\mathrm{X})f_n(\mathrm{X}),$$

并且是具有这个性质的多项式集合中次数最低者.

a) 和 b) 的等价性是显然的. c) 的意思是 d 生成的理想就是 $f_1, \cdots, f_n$ 生成的理想, 而这正是前一节所给的最大公因子的定义. 最后, 注意到 (见定理 2 的证明) 一个理想的生成元是这个理想的非零元中的次数最低者. 由此得到定理 3.

推论 设 $f_1, \cdots, f_n$ 是系数在域 K 内的一个未定元的多项式. 则下列性质是等价的:

a) $f_1, \cdots, f_n$ 的仅有的公因式是 K 的非零元;

b) 存在多项式 $u_i \in \mathrm{K}$, 使得

$$u_1(\mathrm{X})f_1(\mathrm{X}) + \cdots + u_n(\mathrm{X})f_n(\mathrm{X}) = 1.$$

此外, 这个推论还是 §31 定理 1 的特殊情形. 当它成立时, 自然就说多项式 f_i **是互素的**.

不言而喻, 还可以定义 $f_1, \cdots, f_n$ 的**最小公倍式**, 这是 f_i 的具有下列性质的公倍式 m, f_i 的其他非零公倍式都是 m 的倍式; 或者说, 这是 f_i 的公倍式中的次数最小者.

§31 的定理 2, 3, 4 逐字逐句应用到环 K[X], 在这里重述毫无必要.

K[X] 的素元一般称为系数在域 K 内的一个未定元的**不可约多项式**. 所有的多项式 $f \in \mathrm{K}[\mathrm{X}]$ 都可以本质上以一种方式写成不可约多项式的乘积. 这由前一节的定理 5 和 7 得到.



注 1 通过不可约多项式乘以其首项系数的逆元 (这样就把它换为一个 §31 第 6 小节意义下的相伴的多项式), 就会发现可以局限于对首项系数为 1 的不可约多项式实现多项式的不可约因子的分解. 根据引理 1, 两个这样的不可约多项式如果是相伴的必然相等. 换句话说, 如果打算把 §31 第 6 小节的分解 (a) 应用到 K[X], 就需要取首项系数为 1 的不可约多项式的集合作为 P.

注 2 设 p_1 和 p_2 是系数在域 K 内的一个未定元的多项式. 为了计算它们的最大公因式, 可以通过辗转相除法进行: 假定 $d^\circ(p_1) > d^\circ(p_2)$, 我们写出

$$p_1 = p_2v_2 + p_3, \quad 其中\ d^\circ(p_3) < d^\circ(p_2),$$
$$p_2 = p_3v_3 + p_3, \quad 其中\ d^\circ(p_4) < d^\circ(p_3),$$

并且如此下去, 由于 p_i 的次数是减小的, 最终达到关系

$$p_{n-1} = p_nv_n.$$

显然 p_1 和 p_2 的最大公因式也是 p_2 和 p_3 的最大公因式, 故也是 p_3 和 p_4 的最大公因式, 等等, 故也是 p_{n-1} 和 p_n 的最大公因式, 这就是 p_n, 即辗转相除法中最后的非零余式.

这个方法还给出 Bezout 定理一个构造性证明, 即清晰地构造两个多项式 u_1 和 u_2, 使得 $u_1p_1 + u_2p_2 = p_n$. 事实上我们有

$$\begin{aligned}p_n &= p_{n-2} - p_{n-1}v_{n-1}\\ &= (p_{n-4} - p_{n-3}v_{n-3}) - (p_{n-3} - p_{n-2}v_{n-2})v_{n-1}\\ &= p_{n-4} - p_{n-3}(v_{n-3} + v_{n-1}) + p_{n-2}v_{n-2}v_{n-1}\\ &= p_{n-6} - p_{n-5}v_{n-5} - (p_{n-5} - p_{n-4}v_{n-4})(v_{n-3}\\ &\quad +v_{n-1}) + (p_{n-4} - p_{n3}v_{n-3})v_{n-2}v_{n-1},\end{aligned}$$

而继续往下计算就得到我们要找的关系. 参见习题 3.

4. 应用到有理分式

设 K 是一个交换域, 而

$$f(X) = \frac{p(X)}{q(X)}$$

是一个系数在 K 内的一个未定元的有理分式. 写出

$$q(X) = q_1(X)^{r_1} \cdots q_n(X)^{r_n},$$

其中 q_i 是两两不成比例的不可约多项式. §31 的定理 8 指出存在多项式 $p_i(X)(1 \leqslant i \leqslant n)$, 使得有

$$f(X) = \frac{p_1(X)}{q_1(X)^{r_1}} + \cdots + \frac{p_n(X)}{q_n(X)^{r_n}}. \tag{12}$$

我们有下列引理(*)

(*) 我们注意到这个引理类似于初等数论中对于整数 $q \geqslant 2$, 可以有 “以 q 为底的记数法”.

引理 3　设 p 和 q 是系数在一个交换环 K 内的一个未定元的多项式. 假定 q 是归一的, 则存在几乎全部为零的多项式 $h_i \geqslant 0 (i \geqslant 0)$, 使得

$$p = h_0 + h_1 q + h_2 q^2 + \cdots \quad 对于所有 \ i, d^\circ(h_i) < d^\circ(q).$$

因为借助定理 1 写出

$$p = h_0 + p_1 q, \quad d^\circ(h_0) < d^\circ(q),$$

然后有

$$p_1 = h_1 + p_2 q, \quad d^\circ(h_1) < d^\circ(q),$$

如此继续下去, 多项式 p_i 的次数严格递减, 以致对于充分大的 i 有 $p_i = 0$, 把所得到的关系代入前一个, 显然就得到引理.

引理 3 指出对于所有整数 $r > 0$, 可以写出

$$\frac{p}{q^r} = \frac{h_0}{q^r} + \frac{h_1}{q^{r-1}} + \cdots + \frac{h_{r-1}}{q} + p_0,$$

其中的 $h_0, \cdots, h_{r-1}, p_0$ 都是多项式, 并且

$$d^\circ(h_i) < d^\circ(q) \quad 对于 \ 0 \leqslant i \leqslant r-1.$$

把这个结果应用到 (12) 右端出现的每一个分式, 并且合并每一个分式中的项 p_0, 就得到形式为

$$f(\mathrm{X}) = g(\mathrm{X}) + \sum_{i=0}^{n} \sum_{0 \leqslant r \leqslant r_i} \frac{h_{ir}(\mathrm{X})}{q_i(\mathrm{X})^r} \tag{13}$$

的关系, 其中的 g 和 h_{ir} 都是多项式, 并且对于任意 i 和 r

$$d^\circ(h_{ir}) < d^\circ(q_i). \tag{14}$$

公式 (13) 称为给定的有理分式 f 的在**域 K 上的部分分式分解**, “部分分式” 是 $h_{ir}(\mathrm{X})/q_i(\mathrm{X})$. 后面将看到如果 $\mathrm{K} = \mathbf{C}$, 那么 h_{ir} 必然是常数, 而如果 $\mathrm{K} = \mathbf{R}$, 则它们的次数至多是 1.

例 2　取

$$f(\mathrm{X}) = \frac{1}{\mathrm{X}^2(\mathrm{X}-1)^3}.$$

多项式 X 和 X − 1 是一次的, 因此是不可约的. 我们有

$$(\mathrm{X}-1)^3 = \mathrm{X}^3 - 3\mathrm{X}^2 + 3\mathrm{X} - 1 = \mathrm{X}^2(\mathrm{X}-3) + 3\mathrm{X} - 1,$$

继而

$$\mathrm{X}^2 = (3\mathrm{X}-1)\frac{\mathrm{X}}{3} + \frac{\mathrm{X}}{3},$$

再有

$$3X-1=\frac{X}{3}\cdot 9-1,$$

因此 (以上注 2)

$$\begin{aligned}1&=\frac{X}{3}\cdot 9-(3X-1)=9X^2-(3X+1)(3X-1)\\&=9X^2-(3X+1)[(X-1)^3-X^2(X-3)],\end{aligned}$$

于是对于 X^2 和 $(X-1)^3$ 的 Bezout 定理写成

$$1=(3X^2-8X+6)X^2-(3X+1)(X-1)^3.$$

因此有

$$\begin{aligned}f(X)&=\frac{(3X^2-8X+6)X^2-(3X+1)(X-1)^3}{X^2(X-1)^3}\\&=\frac{3X^2-8X+6}{(X-1)^3}-\frac{3X+1}{X^2};\end{aligned}$$

而

$$\begin{aligned}3X^2-8X+6&=(X-1)(3X-5)+1=(X-1)[3(X-1)-2]+1\\&=3(X-1)^2-2(X-1)+1;\end{aligned}$$

终于得到

$$f(X)=\frac{3}{X-1}-\frac{2}{(X-1)^2}+\frac{1}{(X-1)^3}-\frac{1}{X^2}-\frac{3}{X}.$$

这就是所要找的 f 的部分分式分解.

§32 习题

1. 求下列除法的商式和余式:

$$2X^4-3X^3+4X^2-5X+6 \quad 除以 \quad X^2-3X+1,$$
$$4X^3+X^2 \quad 除以 \quad X+1+i,$$
$$X^4-2X^3+4X^2-6X+8 \quad 除以 \quad X-1.$$

2. 求下列多项式的最大公因式:

a) $X^6-7X^4+8X^3-7X+7$ 和 $3X^5-7X^3+3X^2-7$,

b) $X^5+X^4-X^3-3X^2-3X-1$ 和 $X^4-2X^3-X^2-2X-1$,

c) $X^6+X^5-4X^4-2X^3-X^2+X+1, X^5+X^3-X^2-1$ 和 X^4-2X^3-X+2.

3. 对于下面给出的每一对多项式 p 和 q, 求多项式 u 和 v, 使得 $up+vq$ 是 p 和 q 的一个最大公因式.

a) $X^5+3X^4+X^3+X^2+3X+1$ 和 X^4+2X^3+X+2,

b) $3X^5 + 5X^4 - 16X^3 - 6X^2 - 5X - 6$ 和 $3X^4 - 4X^3 - X^2 - X - 6$,

c) $X^5 + 5X^4 + 9X^3 + 7X^2 + 5X + 3$ 和 $X^4 + 2X^3 + 2X^2 + X + 1$,

d) X^4 和 $(1 - X)^4$.

4. 求次数尽可能低的多项式, 使它除以 $X^4 - 2X^3 - 2X^2 + 10X - 7$ 的余式为 $X^2 + X + 1$, 而除以 $X^4 - 2X^3 - 3X^2 + 13X - 10$ 的余式为 $2X^2 - 3$.

¶5. 证明多项式

$$X^m - 1 \quad 和 \quad X^n - 1$$

的最大公因式是 $X^d - 1$, 其中的 d 是整数 m 和 n 的最大公因数.

6. 设 p 和 q 是两个一个未定元的多项式. 如果 $p(X^3) + Xq(X^3)$ 是被 $X^2 + X + 1$ 整除的, 则 $p(1) = q(1) = 0$.

7. 设

$$f(X) = \frac{p(X)}{q(X)}$$

是系数在一个域 K 内的一个未定元的有理分式, 而 a 是其分母 q 的一个单根, 这就使得极点 a 在 f 的部分分式分解中的贡献对于一个 $A \in K$ 是

$$\frac{A}{X - a}.$$

证明有

$$A = \frac{p(a)}{q'(a)}.$$

[写出

$$\frac{p(X)}{q(X)} = \frac{A}{X - a} + \frac{r(X)}{s(X)},$$

其中 $r(a) \neq 0, s(a) \neq 0$, 通分, 求导, 在所得的结果中令 $X = a$. 为了证明这个结果, 在分析教程中一般使用的 "求极限" 过程不能推广到任意域.]

8. 把下列有理分式分解成部分分式 (在 **C** 上, 再在 **R** 上) :

$$\frac{X^2 + 1}{X(X^2 - 1)}, \quad \frac{2}{(X - 1)(X - 2)(X - 3)}, \quad \frac{X^5 - X^3 - X^2}{X^2 - 1}, \quad \frac{4X^3}{(X^2 + 1)^2},$$
$$\frac{X^6 - X^2 + 1}{(X - 1)^3}, \quad \frac{3X^2 + 3}{X^3 - 3X - 2}, \quad \frac{X^5}{(X^4 - 1)^2}.$$

那些认为这些例子不够充分的读者可以容易地构造这样的例子: 方法在于, 任意选择两个多项式 (如果想得到明晰的结果, 只需分母的根是显而易见的或无论如何是可以计算的).

¶9. 设 L 是一个交换域, K 是 L 的一个子域, 而 x 是在 K 上是代数的 L 的一个元素. 设 I 是使得 $f(x) = 0$ 的多项式 $f \in K[X]$ 的集合. 证明这是 $K[X]$ 的一个理想. 由此推出存在唯一的一个系数在 K 内的多项式

$$f(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0, \tag{*}$$

使得 $f(x) = 0$, 并且所有满足 $g(x) = 0$ 的多项式都是 f 的一个倍式. 证明

$$x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0 \tag{**}$$

是 x 所满足的最小次数 (系数在 K 内的) 的方程. 称 (*) 是 x 在 K 上的**极小多项式**, 而 (**) 是 x 在 K 上的**最小方程**, 次数 n 是 x 在 K 上的**次数**. 证明, 看作 K 上的向量空间, 子域 K[X] 是 n 维的, 并且元素

$$1, x, \cdots, x^{n-1}$$

组成它的一个基. 换句话说, 我们采用 §26 中习题 5 的记号有

$$[\mathrm{K}[x] : \mathrm{K}] = n.$$

在前面的叙述中取 $\mathrm{L} = \mathbf{C}, \mathrm{K} = \mathbf{Q}$. 求 L 的下列元素的最小方程:

$$\sqrt{2} + \sqrt{3}; \quad \sqrt[3]{2} + \sqrt{5}.$$

¶10. 设 K 是一个交换域.

a) 设 L 是 K 的一个扩张, 而 x 是 K 上代数的 L 的一个元素. 证明 x 在 K 上的极小多项式是 (K 上) 不可约的.

b) 设 f 是系数在 K 内的一个未定元的不可约多项式, 其首项系数是 1. 设 x 是 f 在 K 的一个扩张内的一个根. 证明 f 是 x 在 K 上的极小多项式.

c) f 是前面问题中那样的多项式, 设 x 和 y 是 f 在 K 的一个扩张 L 内的根. 证明存在唯一一个从域 K(x) 到域 K(y) 上满足

$$j(x) = y, \quad j(a) = a \quad \text{对于所有 } a \in \mathrm{K}$$

的同构 j.

d) 进而假设 K 是一个环 A 的分式域. 沿用问题 c) 的记号, 证明如果 x 在 A 上是整元 (§34, 习题 41), 则 y 也是.

¶¶e) 设 A 是一个交换整环, K 是它的分式域, 而 L 是 K 的一个扩张. 设 $x \in \mathrm{L}$ 在 A 上是整元; 证明 x 在 K 上的极小多项式 f 的系数在 A 上是整元 (把 L 嵌入到一个代数闭域内, 注意到 f 的所有根在 A 上是整元, 并且应用 §33, 第 6 小节). 由此推出, 如果 A 是整封闭的 (即在 A 上是整元的 K 的所有元素是在 A 内的), 则 f 的系数在 A 内.

¶¶f) 假定前面的 L 是 K 的有限扩张 (§26, 习题 4). 证明, 如果 $x \in \mathrm{L}$ 在 A 上是整元, 并且 A 是整封闭的, 则有

$$\mathrm{Tr}_{\mathrm{L/K}}(x) \in \mathrm{A}, \quad \mathrm{N}_{\mathrm{L/K}}(x) \in \mathrm{A}$$

(利用 §26, 习题 5).

¶¶11. 设 L 是一个交换域, K 是 L 的一个子域, 而 x 在 K 上代数的 L 的一个元素. 称 x 在 K 上是**可分的**, 如果它是其在 K 上的极小多项式 f 的单根.

a) x 在 K 上是可分的, 必须并且只需 x 是至少一个系数在 K 内的代数方程的单根.

b) 如果 x 在 K 上不是可分的, 则有 $f'(a) = 0$, 反之亦真.

c) 如果 K 是特征为 0 的, 则在 K 上所有代数的 x 在 K 上是可分的, 并且系数在 K 内的所有不可约多项式的所有根皆为单根.

d) 设 K 的特征 $p \neq 0$, 则对于在 K 上代数的所有的 $x \in \mathrm{L}$ 存在一个整数 $n \geqslant 0$, 使得 x^{p^n} 在 K 上是可分的.

e) 设 L 是 K 的有限代数扩张 (§26, 习题 4). L 在 K 上是可分的 (§26, 习题 4, h)), 必须并且只需所有的 $x \in \mathrm{L}$ 在 K 上是可分的.

¶12. 如果不是常元的多项式 $f \in \mathbf{Z}[\mathrm{X}]$ 在环 $\mathbf{Q}[\mathrm{X}]$ 内不是不可约的, 则可以用非平凡的方式把它分解为整系数多项式的乘积 (利用 Gauss 引理, §27, 习题 13).

13. 以下多项式

$$\mathrm{X}^2+\mathrm{X}+1,\quad \mathrm{X}^4+\mathrm{X}+1,\quad \mathrm{X}^2+\mathrm{X}^2+1,\quad \mathrm{X}^3+7\mathrm{X}+7,\quad \mathrm{X}^5+3\mathrm{X}+2$$

哪些在 $\mathbf{Q}$ 上是不可约的?

¶14. 设 $f(\mathrm{X}) = a_0 + a_1\mathrm{X} + \cdots + a_n\mathrm{X}^n$ 是系数在 $\mathbf{Z}$ 内的一个多项式. 假定某个素数 p 整除 $a_0, \cdots, a_{n-1}$, 不整除 a_n, 并且 a_0 不被 p^2 整除. 证明 f 在 $\mathbf{Q}$ 上是不可约的 (Eisenstein 不可约判别法; 利用习题 12).

¶15. (其元素为多项式的矩阵的初等因子) 设 k 是一个交换域, 而 $\mathrm{K} = k[\mathrm{X}]$ 是系数在 k 内的一个未定元的多项式环.

a) 证明 $\mathrm{GL}(n, \mathrm{K})$ 是这样的矩阵 $\mathrm{U} \in \mathrm{M}_n(\mathrm{K})$ 的集合, 其行列式是域 k 的一个非零元素 (因此这样的矩阵的行列式是 "常" 元).

b) 在以下内容中我们对于元素在 K 内的矩阵使用 §31, 习题 15 的初等变换概念. 设 A 是系数在 K 内的非零矩阵, 并且 $a_{ij}(\mathrm{X})$ 是 A 的第 i 行第 j 列相交处的非零元素. 证明借助有限次初等变换可以把第 i 行的元素或第 j 列的元素换成它们除以 $a_{ij}(\mathrm{X})$ 的余式.

c) 设 $d_1(\mathrm{X})$ 是由 A 经过一系列初等变换所得到的所有矩阵的所有非零元素中最低次数的首项系数为 1 的多项式. 证明通过一系列初等变换可以从 A 得到一个形式为

$$\begin{pmatrix} d_1 & 0 \\ 0 & \mathrm{A}_1 \end{pmatrix}$$

的矩阵, 其中的 A_1 比 A 少一行一列, 并且作为其元素的所有多项式都被 d_1 整除. 证明 d_1 是 A 的非零元素的最大公因式 (因此知道了 A, 就完全确定了 d_1).

d) 证明通过一系列初等变换可以从 A 得到如下形式的矩阵:

$$\begin{pmatrix} d_1(\mathrm{X}) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & d_2(\mathrm{X}) & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_r(\mathrm{X}) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix},$$

其中的 $d_i(\mathrm{X})$ 是非零多项式, 并且每一个整除后面一个. 借助 §31 的习题 16 证明, 对于不超过 A 的行数和列数的所有整数 i, 多项式 $d_1(\mathrm{X}) \cdots d_i(\mathrm{X})$ (当 $i \geqslant r+1$ 时令 $d_i = 0$) 是 A 的 i 阶子式的最大公因式. 因此推知整数 r 等于 A 的秩, 并且对于 $1 \leqslant i \leqslant r$, 多项式 $d_i(\mathrm{X})$ 完全由 A 确定, 如果要求它们的首项系数等于 1.

e) 称 $d_1(\mathrm{X}), \cdots, d_r(\mathrm{X})$ 为矩阵 A 的**不变因子**, 而商 $d_i/d_{i-1}(d_0 = 1, 1 \leqslant i \leqslant r)$ 称为 A 的**初等因子**. 最后, 称元素在 $\mathrm{K} = k[\mathrm{X}]$ 内的 p 行 q 列的两个矩阵 A 和 B 是**等价的**, 如果存在矩阵

$$\mathrm{U} = \mathrm{GL}(p, \mathrm{K}) \quad 和 \quad \mathrm{V} = \mathrm{GL}(q, \mathrm{K}),$$

使得 B = UAV. 证明 A 和 B 是等价的, 必须并且只需 A 和 B 有同样的秩, 和同样的不变因子, 并且可以通过一系列初等变换把一个变换为另一个.

¶¶f) 设 M 是同构于 K^p 的一个 K-模, 而 M′ 是 M 的一个子模. 证明存在 M 的一个基 $(a_1, \cdots, a_p)$ 和多项式 $d_1, \cdots, d_p \in \mathrm{K}$, 使得 d_i 整除 d_{i+1}, 并且 M′ 是由 $d_1a_1, \cdots, d_pa_p$ 生成的 (没有禁止某些 d_i 是零). [应用问题 d) 到 M 的以 M′ 为像的一个自同态关于 M 的一个基的矩阵.]

¶¶g) 设 E 是一个有限生成的 K-模 (不必是自由的). 证明 E 同构于一个形式为 K^s 的模和形式为 $\mathrm{K}/d_1\mathrm{K}, \cdots, \mathrm{K}/d_r\mathrm{K}$ 的模的直积, 其中的 $d_1, \cdots, d_r$ 是非零多项式, 当 $1 \leqslant i \leqslant r-1$ 时 d_i 整除 d_{i+1}. [称 $d_1, \cdots, d_r$ 是 K-模 E 的**不变因子**; 整数 s 是在 §29 习题 11, e) 的意义下的**秩**.] 证明, 当且仅当两个有限生成的 K-模的秩和不变因子都相等, 则它们同构.

¶¶h) 从 §31 习题 8, 9, 10 和 11 以及 K 是主理想整环的事实推出上面的结果. [这个习题的结果, 对于元素在一个域上的一个未定元的多项式环内的矩阵建立了与元素在 **Z** 内的矩阵的初等因子理论 (比如 §31, 习题 17) 类似的理论, 有重要的应用, 尤其是对于常系数任意阶的线性微分方程组; 在参考文献中列举的一些著作 (尤其是 Albert, Gelfand, Schreier-Sperner 等人的) 中可以找到这些结果的精美陈述, 但是这些结果的真正诠释显然是主理想整环上的有限生成模理论.]

在下面的习题 16 至 21(*) 中, 要求化给定的矩阵成习题 15, d) 的典范形式, 并且计算初等因子 (基础域是 **C**).

16. $\begin{pmatrix} \mathrm{X} & 1 \\ 0 & \mathrm{X} \end{pmatrix}$.

17. $\begin{pmatrix} \mathrm{X}^2-1 & \mathrm{X}+1 \\ \mathrm{X}+1 & \mathrm{X}^2+2\mathrm{X}+1 \end{pmatrix}$.

18. $\begin{pmatrix} 1-\mathrm{X} & \mathrm{X}^2 & \mathrm{X} \\ \mathrm{X} & \mathrm{X} & -\mathrm{X} \\ 1+\mathrm{X}^2 & \mathrm{X}^2 & -\mathrm{X}^2 \end{pmatrix}$.

19. $\begin{pmatrix} \mathrm{X} & 1 & 0 & 0 \\ 0 & \mathrm{X} & 1 & 0 \\ 0 & 0 & \mathrm{X} & 1 \\ 0 & 0 & 0 & \mathrm{X} \end{pmatrix}$.

20. $\begin{pmatrix} \mathrm{X} & -1 & 0 & 0 & 0 \\ 0 & \mathrm{X} & -1 & 0 & 0 \\ 0 & 0 & \mathrm{X} & -1 & 0 \\ 0 & 0 & 0 & \mathrm{X} & -1 \\ 0 & 0 & 0 & 4 & 5+\mathrm{X} \end{pmatrix}$.

21. $\begin{pmatrix} \mathrm{X} & 1 & 1 & \cdots & 1 \\ 0 & \mathrm{X} & 1 & \cdots & 1 \\ 0 & 0 & \mathrm{X} & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \mathrm{X} \end{pmatrix}$.

在下面的习题 22 和 23 中, 要求找到其行列式为非零常元而元素为多项式的矩阵 U 和 V, 使得当用 A 表示给定的矩阵时, UAV 取习题 15 的典范形式.

22. $\begin{pmatrix} \mathrm{X}^4+4\mathrm{X}^3+4\mathrm{X}^2+\mathrm{X}+2 & \mathrm{X}^3+4\mathrm{X}^2+4\mathrm{X} \\ \mathrm{X}^4+5\mathrm{X}^3+8\mathrm{X}^2+5\mathrm{X}+2 & \mathrm{X}^3+5\mathrm{X}^2+8\mathrm{X}+4 \end{pmatrix}$.

(*) 习题 16 至 26 摘自 Proskurjakov 的习题集, 在那里读者可以发现许多其他的类似的习题.

23. $$\begin{pmatrix} X^4+3X^3-5X^2+X+1 & 2X^4+3X^3-5X^2+X-1 & 2X^4+2X^3-4X^2 \\ X^4-X^3+1 & 2X^4-X^3-X^2 & 2X^4-2X^3 \\ X^4+2X^3-4X^2+X+1 & 2X^4+2X^3-4X^2+X-1 & 2X^4+X^3-3X^2 \end{pmatrix}.$$

24. 验证元素在 $\mathbf{C}[X]$ 的两个矩阵是等价的:

$$\begin{pmatrix} X^3+6X^2+6X+5 & X^3+4X^2+4X+3 \\ X^3+3X^2+3X+2 & X^3+2X^2+2X+1 \\ 2X^3+3X^2+3X+1 & 2X^3+2X^2+2X \end{pmatrix} = A,$$

$$\begin{pmatrix} X^3+X^2+X & 2X^3+X^2+X-1 \\ 3X^3+2X^2+2X-1 & 6X^3+2X^2+2X-4 \\ X^3-X^2-X-2 & 2X^3-X^2-X-3 \end{pmatrix} = B$$

(计算矩阵 U, V, 使得 B=UAV).

25. 计算矩阵

$$\begin{pmatrix} X^3+X^2-X+3 & X^3-X^2+X & 2X^3+X^2-X+4 & X^3+X^2-X+2 \\ X^3+3X^2-3X+6 & X^3-3X^2+3X-2 & 2X^3+3X^2-3X+7 & X^3+3X^2-3X+4 \\ X^3+2X^2-2X+4 & X^3-2X^2+2X-1 & 2X^3+2X^2-2X+5 & X^3+2X^2-2X+3 \\ 2X^3+X^2-X+5 & 2X^3-X^2+X+1 & 4X^3+X^2-X+7 & 2X^3+X^2-X+3 \end{pmatrix}$$

的不变因子.

26. 计算矩阵

$$\begin{pmatrix} X^4+1 & X^7-X^4+X^3-1 & X^4-4X^3+4X-5 \\ 2X^4+3 & 2X^7-2X^4+4X^3-2 & 3X^4-10X^3+X^2+10X-14 \\ X^4+2 & X^7-X^4+2X^3-2 & 2X^4-6X^3+X^2+6X-9 \end{pmatrix}$$

的初等因子, 取基础环为 $\mathbf{Q}[X]$, 或 $\mathbf{R}[X]$, 或 $\mathbf{C}[X]$.

¶27. 我们打算证明, 如果 K 是一个交换的 Noether 环, 则多项式环 K[X] 是 Noether 环. 用 I 表示 K[X] 的一个理想.

a) 对于所有整数 $n \geqslant 0$, 设 $J_n \subset K$ 是由 0 和满足下列条件的 $a \in K$ 组成的集合: 存在一个首项系数为 a 的 n 次多项式 $f \in I$. 证明 J_n 组成 K 的理想的一个递增序列. 因此推知对于某个整数 r 有

$$J_r = J_{r+1} = \cdots$$

b) 对于所有使得 $0 \leqslant i \leqslant r$ 的整数 i, 在 I 内取有限个 i 次多项式 $f_{ij} (1 \leqslant j \leqslant n_i)$, 它们的首项系数 a_{ij} 生成理想 J_i. 证明, 对于所有 $f \in I$, 存在多项式 $q_{ij} \in K[X]$, 使得

$$f = \sum_{\substack{1\leqslant j\leqslant n_i \\ 0\leqslant i\leqslant r}} q_{ij} f_{ij} + g, \quad 其中\ d^\circ(g) < d^\circ(f).$$

c) 用关于 f 的次数的归纳法推出 $n_0 + \cdots + n_r$ 个多项式 f_{ij} 生成理想 I.

d) 从前面的结果推出, 如果交换环 L 包含一个 Noether 子环 K 和有限个元素 $x_1, \cdots, x_n$, 使得 $L = K[x_1, \cdots, x_n]$, 那么 L 是 Noether 环. (注意到 L 是系数在 K 内的多项式环的一个商.)

¶28. 设 K 是一个有限交换域. 称 K^n 的一个子集 V 是一个代数流形, 如果存在有限个多项式 $p_1, \cdots, p_r \in K[X_1, \cdots, X_n]$, 使得 V 是使 $p_1(x) = \cdots = p_r(x) = 0$ 的 $x \in K^n$ 的集合; 称 K^n 的一个子集 A 是 Zariski 开集, 如果补集 $K^n - A$ 是一个代数流形 (§§27, 28, 习题 1). 利用 $K[X_1, \cdots, X_n]$ 是 Noether 环这一事实证明下列结果:

a) K^n 内的一族 (有限或无限) 代数流形的交集是 K^n 内的代数流形. Zariski 开集的所有并集是 Zariski 开集.

b) K^n 内的所有代数流形的递减序列是稳定的, 所有 Zariski 开集的递增序列是稳定的.

此外证明有

c) K^n 内的代数流形的一个有限族的并集还是 K^n 内一个代数流形, Zariski 开集的有限族的交集还是一个 Zariski 开集.

d) 两个非空 Zariski 开集的交集是非空的. 设 U 和 V 是 K^n 内的两个代数流形, 满足条件 $U \neq K^n$ 和 $V \neq K^n$, 则有 $U \cup V \neq K^n$.

(对于 K^n 内的每一个代数流形 V, 我们的兴趣在于, 引入由在所有 $x \in V$ 都取零值的多项式形成的理想 $I(V) \subset K[X_1, \cdots, X_n]$, 以及用理想的术语解释上述结果.)

¶¶29. 设 K 是交换的 Noether 环. 证明形式幂级数环 K[[X]] (§§27, 28, 习题 11) 是 Noether 环. (给定 K[[X]] 的一个理想 I, 对于 $n \geqslant 0$ 考虑由不含有次数 $\leqslant n-1$ 的项的 $f \in I$ 中的 X^n 的系数组成的理想 J_n.)

¶¶¶30. 设 V 是特征为 0 的交换域 K 上的有限维向量空间, 而 G 是 V 的自同构的 r 阶有限群. 用 A 表示 V 上的多项式函数环 (§§27, 28, 习题 17; 或在情形 $V = K^n$ 下的 §28 第 2 小节显然可以归结为这种情形).

给定一个 $s \in G$ 和 V 上的一个多项式函数, 定义从 V 到 K 的一个新的映射

$$f_s(x) = f(s^{-1}(x)) \quad \text{对于所有 } x \in V.$$

如果对于所有的 $s \in G$ 有 $f_s = f$, 则称 f 是群 G 的一个**不变多项式**. 把这些不变多项式的集合记作 $I \subset A$.

a) 证明对于所有 $f \in A$ 和所有 $s \in G$ 有 $f_s \in A$, 并且 I 是 A 的一个子环.

b) 对于所有的多项式函数 $f \in A$, 定义多项式函数

$$f^{\#} = \frac{1}{r} \sum_{s \in G} f_s.$$

证明 $f^{\#}$ 是 G 的不变多项式. 证明

$$(f+g)^{\#} = f^{\#} + g^{\#} \quad \text{对于所有 } f, g \in A,$$

$$f^{\#} = f \quad \text{当且仅当 } f \in I,$$

$$(fg)^{\#} = f^{\#} g \quad \text{对于所有 } f \in A, g \in I.$$

c) 证明如果 f 是 G 的不变多项式, 则其所有的齐次分量也是 G 的不变多项式 (§§27, 28, 习题 17).

d) 设 J 是环 A 的由 I 生成的理想. 利用问题 c) 和 A 是 Noether 环的事实 (习题 14 和 15), 证明在 I 内存在生成 J 的有限个齐次多项式

$$f_1, \cdots, f_p.$$

下面令 $q_i = d^\circ(f_i)$.

e) 对于所有 q 次齐次多项式 $f \in \mathrm{I}$, 存在 $q-q_i$ 次齐次多项式 $u_i \in \mathrm{A}$ (如果 $q-q_i<0$, 则取 $u_i=0$), 使得 $f=\sum f_i u_i$. 证明甚至可以在 I 内取 u_i (写出 $f=f^{\#}$).

f) 通过关于 f 的次数的归纳推理, 由上推出所有 $f \in \mathrm{I}$ 是系数在 K 内的 f_i 的一个多项式, 换句话说, 群 G 的不变多项式组成由有限个元素生成的 K 上的一个环 (Hilbert 的不变多项式定理).

¶¶31. (这个习题的解答假定承认了 §31 习题 21 的结果.) 我们打算证明: 如果 A 是一个唯一因式分解整环, 则环 A[X] 也是唯一因式分解整环.

a) 给定多项式 $f, g \in \mathrm{A[X]}$, 设 A 的不可约元素 p 整除 fg 的所有系数. 证明 p 整除 f 的所有系数, 或整除 g 的所有系数.

b) 称一个多项式 $f \in \mathrm{A[X]}$ 是**本原的**, 如果它的系数的最大公因子是 1. 证明如果 f 和 g 是本原的, 则 fg 也是.

c) 对于所有非零的 $f \in \mathrm{A[X]}$, 用 $c(f)$ 表示它的系数的一个最大公因子. 证明

$$c(fg) = c(f)c(g)$$

(对于唯一因子分解整环的 Gauss 引理).

d) 设 K 是 A 的分式域. 证明如果 $f \in \mathrm{A[X]}$ 在 K[X] 内不是不可约的, 则它在 A[X] 内也不是不可约的.

e) 由此和问题 a) 推出环 A[X] 的不可约元是唯一因子分解整环 A 的不可约元, 以及本原的, 并且在 K[X] 内不可约的非常元的多项式.

f) 由此和 §32 第 3 小节 (应用到 K[X]) 的结果推出 A[X] 的所有元素可以用本质上唯一的方式写成 A[X] 的不可约元乘积的形式, 因此像所宣布的那样 A[X] 是唯一因子分解整环.

g) 证明如果 A 是一个唯一因子分解整环 (例如 A 是一个域, 或 $\mathrm{A} = \mathbf{Z}$), 则环 $\mathrm{A}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$ 是唯一因子分解整环. 作为特殊情形, 所有系数在一个域 K 内的 (n 个未定元的) 多项式可以用本质上唯一的方式分解为系数在一个域 K 内的不可约多项式的乘积 (称系数在 K 内的一个多项式是**不可约的**, 如果它不是常元, 并且它的每一个因子是常元或是正比于 f 的多项式).

h) 证明环 $\mathbf{Z}$[X] (根据前一个问题它是唯一因子分解整环) 不是主理想整环 (考察 2 和 X 生成的理想). 对于 K[X,Y] 解答同样的问题, 其中 K 是一个域.

i) 证明, 对于所有的交换域 K, 多项式 $\mathrm{Y}^2-\mathrm{X}^3$ 在环 K[X,Y] 内是不可约的 (写出 K[X,Y]=A[Y], 其中 A=K[X], 并且利用上面的问题 d)).

j) 证明系数在域 K 内的 n 个未定元的所有有理分式 f, 可以写成形式 $f=p/q$, 其中的 p 和 q 是系数在 K 内的 n 个未定元的多项式, 并且是互素的 (即没有常元之外的公因子), 此外如果不计常因子 p 和 q 是唯一的. 证明 f 的未定点是 K^n 的使得 p 和 q 同时为零的点, 而其极点是使得 $p(x) \neq 0$ 和 $q(x)=0$ 的点 $x \in \mathrm{K}^n$.

k) 设 p 和 q 是两个系数在交换域 K 内的 $n \geqslant 2$ 个未定元的不是常元的多项式. 假定 p 和 q 是互素的. 能否推出存在系数在 K 内的 n 个变量的多项式 u 和 v, 使得

$$up+vq=1?$$

[以环 $\mathbf{C}[X_1, \cdots, X_n]$ 为例, 它是唯一因子分解整环的几何解释如下所述. 设 $f \in \mathbf{C}[X_1, \cdots, X_n]$ 不是常元, 而 V 是 $\mathbf{C}^n$ 的由方程 $f(x) = 0$ 定义的超曲面. 设

$$f(X) = \prod_{i=1}^{s} p_i(X)^{r_i}$$

是 f 的不可约因式的乘积的分解, 并且设 V_i 是超曲面 $p_i(x) = 0$. 那么 V_i 是不可约的 (即不可能以非平凡的方式表示成两个另外的代数流形的并集), 我们有

$$V = V_1 \cup \cdots \cup V_s,$$

并且不计次序 V 的这个不可约曲面的分解是唯一的.

还可以以下列观点看待这个问题. 设 V 是 $\mathbf{C}^n$ 中的一个代数流形, 而 $\mathfrak{a}$ 是由对于所有 $x \in V$ 都有 $f(x) = 0$ 的多项式组成的 $\mathbf{C}[X_1, \cdots, X_n]$ 的理想. 由于 $\mathbf{C}[X_1, \cdots, X_n]$ 是 Noether 环, §18 的习题 9 的 b) 证明 $\mathfrak{a}$ 是 $\mathbf{C}[X_1, \cdots, X_n]$ 的有限个准素理想的交集 (甚至是素理想的交集: 利用 §18 的习题 11, 并且注意到, 鉴于显然的关系

$$\text{在 V 上 } f(x)^q = 0 \text{ 蕴含在 V 上 } f(x) = 0,$$

理想 $\mathfrak{a}$ 等于它的根). 因此我们可以写出

$$\mathfrak{a} = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_s,$$

其中的 $\mathfrak{p}_i$ 是 $\mathfrak{a}$ 的极小素理想, 设 V_i 是由满足关系

$$f(x) = 0 \quad \text{对于所有 } f \in \mathfrak{p}_i$$

的 x 组成的 $\mathbf{C}^n$ 的代数流形 (如果愿意, V_i 是由有限个方程定义的: 取 $\mathfrak{p}_i$ 的生成元的集合). 由于 $\mathfrak{p}_i$ 是素理想, 每个 V_i 是不可约的, 并且有

$$V = V_1 \cup \cdots \cup V_s;$$

称 V_i 是 V 的**不可约分支**. 这些交代了之后, 环 $\mathbf{C}[X_1, \cdots, X_n]$ 是唯一因式分解整环这一事实指出, 如果 V 是*超曲面* (即由唯一的方程定义), 那么它的不可约分支也是超曲面; 或说成: 如果一个不可约流形 W 包含在一个超曲面 V 内, 则存在一个不可约*超曲面* W′, 使得 $W \subset W' \subset V$ —— 几何上是 "显然的" 结果.

作为唯一因式分解整环的其他的另一个重要例子, 我们举出 (Weierstrass) 的 n 个复变量的*收敛* (即收敛域不退化为 0 的幂级数) 环. 这个环涉及在 $\mathbf{C}^n$ 内的 "解析流形" (由左端是全纯函数的方程定义的 $\mathbf{C}^n$ 的子集) 的 "局部" 研究. 这个环也是 Noether 环].

¶¶ 32. 推广 Eisenstein 不可约判别法 (习题 14) 到唯一因子分解整环.

§33 代数方程的根

1. 根的最大数目

我们已经证明 (§§27, 28, 引理 2) 系数在一个交换整环 K 内的一个未定元的 n 次多项式至多具有 n 个根. 在下一个定理的推论中将会改进这一结果.

定理 1 设 f 是系数在一个交换域 K 内的一个未定元的 n 次多项式. 设 $a_1, \cdots, a_p$ 是 f 在 K 内的所有根, 而 $r_1, \cdots, r_p$ 是它们的重数. 则有

$$f(\mathrm{X}) = (\mathrm{X} - a_1)^{r_1} \cdots (\mathrm{X} - a_p)^{r_p} g(\mathrm{X}), \tag{1}$$

其中的 g 在 K 内没有任何根.

首先根据根的重数的定义, f 被每个多项式

$$(\mathrm{X} - a_1)^{r_1}, \cdots, (\mathrm{X} - a_p)^{r_p} \tag{2}$$

整除. 其次所有形式为 $\mathrm{X} - a$ 是不可约的, 因为如果

$$\mathrm{X} - a = p(\mathrm{X})q(\mathrm{X}),$$

则有 $1 = d^\circ(p) + d^\circ(q)$, 故多项式 p, q 中有一个是 0 次的, 即是环 K[X] 的一个可逆元. 最后对于 $a \neq b$, 多项式 $\mathrm{X} - a$ 和 $\mathrm{X} - b$ 不是相伴的 (即不是成比例的).

由此得到多项式 (2) 是两两互素的 (§31, 引理 6), 因此 (§31, 定理 4) 它们的乘积整除 f, 由此得到存在关系 (1).

关系 (1) 表明 g 在 K 内的所有的根 a 是 f 在 K 内的根, 从而是 $a_1, \cdots, a_p$ 中的一个. 但是比如说 a_1 是 g 的一个根, 多项式 $g(\mathrm{X})$ 将被 $\mathrm{X} - a_1$ 整除, 而 f 将被 $(\mathrm{X} - a_1)^{r_1+1}$ 整除, 这与根的重数相矛盾. 故 g 在 K 内没有任何根. 这就完成了证明.

推论 沿用定理 1 的记号, 则有

$$r_1 + \cdots + r_p \leqslant n = d^\circ(f).$$

这是显然的, 因为有

$$d^\circ(f) = r_1 + \cdots + r_p + d^\circ(g). \tag{3}$$

这个推论表示 f 在 K 内的根的数目不超过 n, 即使 r 重根当作 r 个单根计数也一样.

例如假定 f 是三次多项式; 那么 f 在 K 内的根的数目仅有的可能性如下:

a) f 仅有单根, 那么至多有三个单根;

b) f 有一个二重根, 那么或者 f 仅有一个根, 该根是二重根, 或者一个二重根, 一个单根;

c) f 有一个三重根, 那么这是唯一的三重根, 并且没有其他的根.

事实上, 这些可能性的表达还不完整, 还要考虑以下事实: 如果 f 具有两个不同单根, 或一个二重根, 那么 f 必定在 K 内有另一个根; 事实上, 如果有

$$f(\mathrm{X}) = (\mathrm{X} - a)(\mathrm{X} - b)g(\mathrm{X}),$$

多项式 $g(X)$ 必然是一次的, 故对于一个适当的 c 正比于 $X - c$, 从而 c 是 f 的根.

换句话说, 对于一个三次多项式有以下几种可能性发生:

没有任何根,

一个单根,

三个单根,

一个二重根和一个单根,

一个三重根.

容易给出每一种情形的例子. 对于第一种情形, 取(*)

$$K = \mathbf{Q} \quad 和 \quad f(X) = X^3 - 2;$$

对于其余情形, 只需取 $K = \mathbf{R}$ 和多项式

$$(X-1)(X^2+1), \quad (X-1)(X-2)(X-3), \quad (X-1)(X-2)^2, \quad (X-1)^3.$$

再回到任意域 K 和任意多项式的一般情形. 在定理 1 的分解中, 可能会遇到 g 是常元即 0 次多项式的情形, 换言之, 有

$$f(X) = c(X-a_1)^{r_1} \cdots (X-a_p)^{r_p},$$

其中 c 必然是 f 的首项系数. 或同样的, f 可以写成系数在 K 内的一次多项式的乘积. 如果是这样, 就说 f 的**所有根都在 K 内**. 采用这个术语的理由如下: 如果 L 是 K 的一个扩张, 那么 f 在 L 内的根, 即使得

$$c(x-a_1)^{r_1} \cdots (x-a_p)^{r_p} = 0$$

的 $x \in L$ 恰好是 $a_1, \cdots, a_p$, 即 f 在 K 内的根. 后面将看到, 反之, 如果 f 的根不是全部在 K 内, 就存在 K 的一个扩张 L (例如包含 K 的代数闭域, 参见下一小节) 和 f 在 L 内的但是不在 K 内的根.

为了 f 的所有根都在 K 内, 显然必须并且只需 f 在 K 内的根的重数 $r_1, \cdots, r_p$ 满足关系

$$r_1 + \cdots + r_p = d^\circ(f),$$

即 f 在 K 内的根的数目 (每个 r 重根计算 r 次) 等于 f 的次数.

最后注意如果 *f 的所有根都在 K 内, 则 f 的所有因式 g 也如此*. 事实上, g 是不可约多项式的乘积, 这些不可约多项式整除 g, 必然也整除 f; 而 f 是一次多项式

(*) 不可能在第一个例子中取 $K = \mathbf{R}$, 因为在分析中证明了所有实系数的奇次多项式至少有一个实根 (简略地回忆其证明: 首先指出对于变量的十分大的值, 多项式的性态与其最高次项相同, 由此与该项同符号, 由此得到一个奇次多项式在 $\mathbf{R}$ 上不能保持同样符号, 那么对于连续函数的中值定理指出一个这样的多项式必定在某一个点变为零).

此类结果虽然涉及多项式, 实际上属于分析. 这就是为什么本书不予陈述的理由.

的乘积, 这些一次多项式是不可约的, f 的仅有的不可约因式就是这些一次因式, 因此 g 的不可约因式本身也是一次的, 因此 g 是一次因式的乘积, 这就确立了我们的断言. (还可以应用 §31 的第 7 小节的引理 4 到 f 和 g, 该引理告诉我们如何从 f 的不可约因式分解得到其所有因式.)

2. 代数闭域

设 K 是一个交换域, 而 f 是一个系数在 K 内的一个未定元的多项式, 并且其次数 $n \geqslant 1$. 按照前面所讲的, f 在 K 内至多具有 n 个根. 但是我们还从来没有宣布过任何断定 f 在 K 内的根的*存在性*的定理 (理由不言而喻, 像多项式 X^2+1 在 $\mathbf{R}$ 内就没有任何根).

我们称一个交换域 K 是**代数闭域**, 如果所有系数在 K 内的一个未定元的次数 $\geqslant 1$ 的多项式至少在 K 内有一个根. 关于这类域的存在性, 两个基本结果如下:

定理 2 (d'Alembert-Gauss) *复数域是代数闭域.*

定理 3 (Steinitz) *所有交换域可以嵌入到一个代数闭域内.*

定理 2 告诉我们*所有复系数的次数 $n \geqslant 1$ 的代数方程*

$$a_0 + a_1x + \cdots + a_nx^n = 0$$

至少有一个复根. 比起仅仅为了使二次方程有根而发明复数来, 这个结果更令人惊异. 至于 Steinitz 定理, 它的意思是, 对于所有的交换域 K, 可以构造一个代数闭域 L, 使得 L 包含一个与 K 同构的子域. 如果 $K = \mathbf{R}$, 例如就可以取 $L = \mathbf{C}$ (在一般情形, 从 K 出发构造 L 要比从 $\mathbf{R}$ 出发构造 $\mathbf{C}$ 复杂得多, 不过借助类似的方法可以实现).

不求助不论哪种方式的属于分析而非代数的考虑, 证明定理 2 是不可能的. 参见在本节习题 25 中的简单证明. 至于定理 3, 它的证明远远超出了本著作的范围, 参见习题 20.

自然会问是否一个*未定元*的代数方程根的存在性推导出*多个未定元*的代数方程组的类似性质. 这个问题的回答是归功于 Hilbert 的最著名定理之一:

定理 4 (Hilbert 的零点定理) *设 K 是交换的代数闭域, 而 I 是环 $K[X_1, \cdots, X_n]$ 的一个理想. 则以下性质是等价的:*

a) *至少存在一个点 $x \in K^n$, 使得*

$$f(x) = 0 \quad 对于所有\ x \in I; \tag{4}$$

b) $I \neq K[X_1, \cdots, X_n]$.

既然多项式 1 在 K^n 的任何点不变为零, 那么 a) 蕴含 b) 就是显然的; 但是 b) 蕴含 a) 这个事实的证明却太难而不在这里陈述了, 参见本节的习题 33.

为了了解定理 4 的表述, 我们给定多项式

$$f_1, \cdots, f_p \in \mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n],$$

并且打算在 K^n 内解**代数方程组**

$$f_1(x) = \cdots = f_p(x) = 0. \tag{5}$$

设 I 是 $\mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$ 的由 $f_1, \cdots, f_p$ 生成的理想, 即形式为

$$f = u_1 f_1 + \cdots + u_p f_p$$

的多项式的集合. 显然 (5) 的解和 (4) 的解是相同的, 故说 (5) 至少有一个解就意味着 I 不是整个多项式环, 或归结为同一件事, 即

$$1 \notin \mathrm{I}.$$

换句话说, 有

推论 给定一个交换的代数闭域 K 和系数在 K 内 n 个未定元的多项式

$$f_1, \cdots, f_p \in \mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n],$$

则以下性质是等价的:

a) 代数方程组

$$f_1(x) = \cdots = f_p(x) = 0$$

没有任何解 $x \in \mathrm{K}^n$;

b) 存在多项式

$$u_1, \cdots, u_p \in \mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n],$$

使得

$$u_1 f_1 + \cdots + u_p f_p = 1.$$

注 1 设 K 是一个交换的代数闭域, K^n 的一个子集称为**仿射代数流形**, 如果它是一个代数方程组 $f_1(x) = \cdots = f_p(x) = 0$ 解的集合. 这类仿射代数流形 (及类似对象) 的研究是代数几何的目的. 由仅一个方程 $f(x) = 0$ (其中 f 非常元) 定义的代数流形称为一个**超曲面** (当 $n = 3$ 时一个**曲面**, 或当 $n = 2$ 时**平面曲线**). 因而定理 4 的推论是代数超曲面的交集为空集的一个充分必要条件.



3. 系数在代数闭域内的方程根的数目

在一个代数闭域上, 可以显著改进定理 1:

定理 5 设 f 是一个系数在代数闭域 K 内的一个未定元的多项式, 并且其次数 $n \geqslant 1$. 又设 $a_1, \cdots, a_p$ 是 f 在 K 内的所有不同的根, 而 $r_1, \cdots, r_p$ 是它们的重数. 则有

$$f(\mathrm{X}) = c(\mathrm{X} - a_1)^{r_1} \cdots (\mathrm{X} - a_p)^{r_p}, \tag{6}$$

其中 c 是 f 的首项系数. 此外还有

$$r_1 + \cdots + r_p = n. \tag{7}$$

只需应用定理 1, 由于 g 在 K 内不变为零, 那么根据代数闭域的定义, g 是一个常元, 由此得到所陈述的第一个断言. 第二个断言由计算关系 (6) 右端的次数而得到.

关系 (7) 一般如下表述: 在一个代数闭域内, 一个次数 $n \geqslant 1$ 的代数方程恰好具有 n 个根, 只要认为 r 重根是 r 个不同的根.

例 1 在一个交换域内考虑方程

$$x^n = 1 \tag{8}$$

(它的根是单位元在 K 内的 n **次方根**). 令

$$f(\mathrm{X}) = \mathrm{X}^n - 1, \quad 于是\ f'(\mathrm{X}) = n\mathrm{X}^{n-1},$$

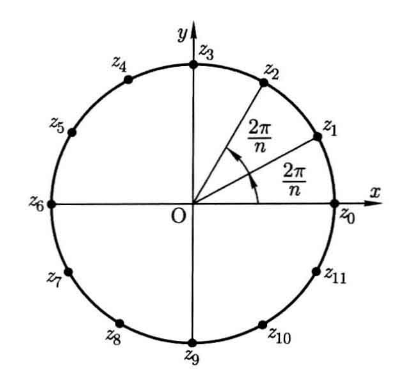
我们发现如果 n 不是 K 的特征的倍数, f' 的仅有的根是 0, 0 显然不是单位元的 n 次方根, 因此 (§30, 定理 6) 在这里的假定下 f 的所有的根都是单根, 如果进一步假设 K 是代数闭域, 那么在 K 内刚好有方程 (8) 的 n 个不同的单根. 在 $\mathrm{K} = \mathbf{C}$ 时正是这种情形.

当 $\mathrm{K} = \mathbf{C}$ 时, 单位的 n 次方根由下式给定:

$$z_k = \cos(2k\pi/n) + i\sin(2k\pi/n) \quad (0 \leqslant k \leqslant n-1),$$

在复平面上它们由单位圆的内接正 n 边形的顶点表示 (见下图), 事实上, de Moive 公式指出

$$z_k^n = \cos(2k\pi n/n) + i\sin(2k\pi n/n) = 1,$$



于是上面的公式表示方程 (8) 的两两不同的 n 个根. 因为这个方程是 n 次的, 不可能再有其他的根.

总是假定 $\mathrm{K} = \mathbf{C}$, 考虑更一般的方程

$$z^n = a,$$

这里 a 是一个给定的非零复数 (它的根称为 a 的 n **次方根**). 写出

$$z = \rho(\cos\theta + i\sin\theta), \quad a = r(\cos\varphi + i\sin\varphi),$$

其中的 r 和 ρ 是正实数, 而 θ 和 φ 是实数, 问题归结为写出

$$\rho^n[\cos(n\theta) + i\sin(n\theta)] = r(\cos\varphi + i\sin\varphi).$$

左端的绝对值是 ρ^n, 辐角是 $n\theta$ (精确到 2π 的倍数: 一个辐角由定义是模 2π 的实数), 右端的绝对值是 r, 而辐角是 φ, 因此方程 $x^n = a$ 等价于方程

$$\rho^n = r, \quad n\theta \equiv \varphi (\mathrm{mod}\ 2\pi).$$

于是它有复根

$$z_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2k\pi}{n} + i\cdot\sin\frac{\varphi + 2k\pi}{n}\right),$$

由于显然有 $z_k = z_{k+n}$, 只需给整数 k 以相继的 n 个值就得到所考虑的方程的所有的根 (两两不同的 n 个数).

注 2 如果域 K 不是代数闭域, 定理 5 显然不成立. 但是在如下情形它仍然为真, 如果考虑 f 在一个包含 K 的代数闭域上的根 (该域的存在性由 Steinitz 定理得到). 比如, 如果 f 是实系数的 n 次多项式, 一般它有 n 个实根的断言是错误的; 但是 f 总是具有 n 个实根或复根是成立的 (根的数目的计算当然要考虑到根的重数). 例如, 多项式 $\mathrm{X}^2 + 1$ 具有两个实根或复根, 即 i 和 $-i$. 同样, 方程



$$x^n = 1$$

在 $\mathbf{R}$ 内仅有一个根 (如果 n 是奇数) 或两个根 (如果 n 是偶数), 而在 $\mathbf{C}$ 内有 n 个根.

4. 系数在代数闭域内的不可约多项式

设 $f(\mathrm{X})$ 是系数在代数闭域内的一个未定元的多项式, 那么 f 是不可约的, 必须并且只需 f 是一次的, 即其形式为

$$f(\mathrm{X}) = a\mathrm{X} + b, \quad a \neq 0.$$

条件显然是充分的 (即使 K 不是代数闭域), 因为 f 的仅有的因式是零次的 (即常元) 或 1 (即正比于 f). 反之, 假定 f 是不可约的, 那么 f 在环 K[X] 内不是可逆的, 故其次数 $n \geqslant 1$ (§32, 引理 2). 由于 K 是代数闭域, f 至少有一个根 $a \in \mathrm{K}$, 但这时 f 是 $\mathrm{X} - a$ 的倍式, 由于 $\mathrm{X} - a$ 是不可约的, 从而 f 正比于 $\mathrm{X} - a$, 我们的断言证明完毕.

当利用 §31 第 6 小节末尾的公式 (9) 实施一个多项式 $f \in \mathrm{K[X]}$ 的不可约因式分解时, 可以取形如

$$\mathrm{X} - a, \quad a \in \mathrm{K}$$

的多项式集合作为 P, 那么 §31 的公式 (9) 显然就化作上面定理 5 的公式 (6).

这个公式构成 f 的在主理想整环 K[X] 内的不可约因式分解, 比如可以用它求两个多项式的最大公因式. 用

$$a_1, \cdots, a_m, \quad b_1, \cdots, b_n$$

表示 f 的不同的根, 用

$$a_1, \cdots, a_m, \quad c_1, \cdots, c_p$$

表示 g 的不同的根, 其中把 f 和 g 的可能有的公共根明显表示了出来 (于是假定对所有 $i, j, b_i \neq c_j$). 写出

$$\begin{aligned} f(\mathrm{X}) &= u(\mathrm{X} - a_1)^{r'_1} \cdots (\mathrm{X} - a_m)^{r'_m} (\mathrm{X} - b_1)^{s_1} \cdots (\mathrm{X} - b_n)^{s_n}, \\ g(\mathrm{X}) &= v(\mathrm{X} - a_1)^{r''_1} \cdots (\mathrm{X} - a_m)^{r''_m} (\mathrm{X} - c_1)^{t_1} \cdots (\mathrm{X} - c_p)^{t_p}, \end{aligned}$$

其中 u 和 v 是非零常元, 由 §31 第 7 小节, 显然 f 和 g 的最大公因式由公式

$$d(\mathrm{X}) = (\mathrm{X} - a_1)^{r_1} \cdots (\mathrm{X} - a_m)^{r_m}, \quad r_i = \min(r'_i, r''_i)$$

给定. 换句话说, 最大公因式的根是 f 和 g 的公共根, 并且如果一个公共根对于 f 是 r' 重的, 而对于 g 是 r'' 重的, 则对于 f 和 g 的最大公因式是 $\min(r'_i, r''_i)$ 重的.

定理 5 在代数闭域的情形改进了 §32 第 4 小节的有理分式的部分分式分解. 事实上, 在这个情形, 不可约多项式 q_i 由

$$q_i(\mathrm{X}) = \mathrm{X} - a_i$$

给定, 并且是一次的. 出现在 §32 公式 (12) 中的多项式 h_{ir} 最高是零次的, 即是常元, 最终发现在代数闭域上, 所有有理分式可以写成形式

$$f(\mathrm{X}) = g(\mathrm{X}) + \sum_{i=1}^{n} \sum_{0 \leqslant r \leqslant r_i} \frac{c_{ir}}{(\mathrm{X} - a_i)^r}, \tag{9}$$

其中的 c_{ir} 是 K 的元素, a_i 是 f 的分母的不同的根, r_i 是这些根的重数, 而 g 是一个多项式 (容易验证它是 f 的分子除以分母的商式).

5. 实系数不可约多项式

第 4 小节的结果应用到复数域 $\mathbf{C}$ 却不能应用到实数域 $\mathbf{R}$, 因为它不是代数封闭的. 虽然这样, 还是有完整的结果:

定理 6 实系数的一个未定元的多项式环 $\mathbf{R}[\mathrm{X}]$ 的不可约元素是多项式

$$a\mathrm{X} + b, \quad 其中\ a \neq 0$$

或多项式

$$a\mathrm{X}^2 + b\mathrm{X} + c, \quad 其中\ b^2 - 4ac < 0.$$

显然对于所有交换域 K, 一次多项式是 K[X] 的不可约元. 在 K 内没有根的二次多项式也是 K[X] 的不可约元, 因为一个这样的多项式的非平凡因式必然是一次的, 即其形式为 $a\mathrm{X} + b$, 而如果 f 被 $a\mathrm{X} + b$ 整除, 那么 $-b/a$ 就是 f 在 K 内的根.

对于 $\mathrm{K} = \mathbf{R}$, 还要确认除上面列举的之外, 再没有其他的不可约多项式.

设 f 是 $\mathbf{R}[\mathrm{X}]$ 的一个不可约元 (从而至少是一次的). 如果 f 在 $\mathbf{R}$ 内具有一个根 a, 那么 f 被 $\mathrm{X} - a$ 整除, 从而正比于 $\mathrm{X} - a$, 因此 f 是一次的, 其逆显然成立.

假定 f 在 $\mathbf{R}$ 内没有任何根. 由于 $\mathbf{C}$ 是代数闭域, 所考虑的多项式

$$f(\mathrm{X}) = a_0 + a_1\mathrm{X} + \cdots + a_n\mathrm{X}^n$$

至少有一个复根

$$w = u + iv \quad (u, v \in \mathbf{R}),$$

但是它必定还有 w 的共轭复数

$$\bar{w} = u - iv$$

也是根, 这是因为系数 a_i 为实数, 我们有

$$f(\bar{w}) = a_0 + a_1\bar{w} + \cdots + a_n\bar{w}^n = \bar{a}_0 + \overline{a_1 w} + \cdots + \overline{a_n w^n} = \overline{f(w)},$$

由此即得我们的断言. 因此在环 $\mathbf{C}[\mathrm{X}]$ 内 f 被 $\mathrm{X} - w$ 和 $\mathrm{X} - \bar{w}$ 整除. 又因为 $w \neq \bar{w}$, 这两个多项式互素, 所以 f 被

$$(\mathrm{X} - w)(\mathrm{X} - \bar{w}) = (\mathrm{X} - u - iv)(\mathrm{X} - u + iv) = (\mathrm{X} - u)^2 + v^2$$

整除, 上式是一个实系数的没有实根的二次多项式. 为了推导出 f 是二次的只需指出 f 不仅在 $\mathbf{C}[\mathrm{X}]$ 被 $(\mathrm{X}-u)^2+v^2$ 整除, 而且在 $\mathbf{R}[\mathrm{X}]$ 内也被 $(\mathrm{X}-u)^2+v^2$ 整除.

我们以一般的方式来考虑一个其扩张为 L 的域 K 和系数在 K 内的两个多项式 f 和 g. 设 q 和 r 是在 K[X] 内 f 除以 g 的商式和余式, 这是系数在 K 内的多项式, 它们满足关系

$$f=gq+r, \quad d^{\circ}(r)<d^{\circ}(g).$$

显然把 f,g,q,r 看作系数在 L 内的多项式这些关系仍然成立, 故知 f 除以 g 的商式和余式在环 K[X] 内和在 L[X] 内是相同的. 特别是如果在 L[X] 内 g 整除 f, 那么在 K[X] 内也如是. 这就完成了定理的证明.

由定理 6 和 §31 定理 5 得到, 所有实系数非零多项式可以写成形式

$$f(\mathrm{X})=u\cdot(\mathrm{X}-a_1)^{r_1}\cdots(\mathrm{X}-a_p)^{r_p}(\mathrm{X}^2+b_1\mathrm{X}+c_1)^{s_1}\cdots(\mathrm{X}^2+b_q\mathrm{X}+c_q)^{s_q},$$

其中的 u 是 f 的首项系数, a_i 是 f 在 $\mathbf{R}$ 内的不同的根, r_i 是它们的重数, 多项式 $\mathrm{X}^2+b_j\mathrm{X}+c_j$ 没有实根, 即其系数满足关系 $b_j^2-4c_j<0$.

还可以如下得到 f 的这个分解. 上面已经从关系

$$f(\bar{w})=\overline{f(w)}$$

得到了 f 的复根的共轭复数仍然是 f 的一个根, 那么 f 就具有偶数个非实根, 设为 $2q$. 用

$$w_j=u_j+iv_j \quad (1\leqslant j\leqslant q)$$

表示虚部为正的根, 那么其余的是它们的共轭复数

$$\bar{w}_j=u_j-iv_j \quad (1\leqslant j\leqslant q).$$

因此在环 $\mathbf{C}[\mathrm{X}]$ 内, 则有公式

$$f(\mathrm{X})=u(\mathrm{X}-a_1)^{r_1}\cdots(\mathrm{X}-a_p)^{r_p}(\mathrm{X}-w_1)^{s_1}\cdots(\mathrm{X}-w_p)^{s_q}(\mathrm{X}-\bar{w}_1)^{s_1'}\cdots(\mathrm{X}-\bar{w}_p)^{s_q'}.$$

而且事实上, 从 §30 定理 7 得知共轭复根的重数是相同的, 合并这个分解里的对应的因式, 我们得到因式

$$[(\mathrm{X}-w_j)(\mathrm{X}-\bar{w}_j)]^{s_j}=[(\mathrm{X}-u_j)^2+v_j^2]^{s_j},$$

由此即得在 $\mathbf{R}[\mathrm{X}]$ 内的不可约因式分解.

像在前一小节一样, 这些结果让我们可以改进在 §32 第 4 小节给出的有理分式的部分分式分解. 事实上, 一般情形的多项式 q_i 在这里有两类: 一类是

$$q_i(\mathrm{X})=\mathrm{X}-a_i,$$

对应分母的实根 (部分分式的分子 h_{ir} 的次数 <1, 即是常数); 另一类是

$$q_j(\mathrm{X}) = \mathrm{X}^2 + b_j\mathrm{X} + c_j, \quad \text{其系数满足 } b_j^2 - 4c_j < 0,$$

它们对应分母的一对共轭复根 (多项式 h_{jr} 的次数 <2, 即有形式 $u_{jr}\mathrm{X} + v_{jr}$, 其中的 u_{jr}, v_{jr} 是常数). 因此, §32 的公式 (13) 在实数域的情形, 呈形式

$$f(\mathrm{X}) = g(\mathrm{X}) + \sum_{i=1}^{m} \sum_{0\leqslant r\leqslant r_i} \frac{c_{ir}}{(\mathrm{X}-a_i)^r} + \sum_{j=1}^{s} \sum_{0\leqslant r\leqslant s_i} \frac{u_{jr}\mathrm{X} + v_{jr}}{(\mathrm{X}^2 + b_j\mathrm{X} + c_j)^r}.$$

这个公式在有理分式的原函数的求法中起关键作用. 为了对于这一事实深信不疑, 读者只需查阅一下 150 年来巴黎综合工科学校入学考试的口试题列表.

6. 方程的根与系数的关系

设

$$f(\mathrm{X}) = u_n\mathrm{X}^n + u_{n-1}\mathrm{X}^{n-1} + \cdots + u_0 \tag{10}$$

是系数在交换环 K 内的次数 $n \geqslant 1$ 的一个多项式, 并且假定 f 具有 n 个根 (考虑这些根的重数), 例如当 K 是代数闭域时就是这样. 用

$$a_1, \cdots, a_n$$

表示这些根, 每个 r 重根写 r 次. 根据定理 1 有关系

$$f(\mathrm{X}) = u_n(\mathrm{X}-a_1)(\mathrm{X}-a_2)\cdots(\mathrm{X}-a_n). \tag{11}$$

令

$$(\mathrm{X}-a_1)(\mathrm{X}-a_2)\cdots(\mathrm{X}-a_n) = \mathrm{X}^n + v_{n-1}\mathrm{X}^{n-1} + \cdots + v_0, \tag{12}$$

利用 §8 第 5 小节的公式

$$\prod_{i\in\mathrm{I}} (x_i + y_i) = \sum_{\mathrm{F}\subset\mathrm{I}} x_{\mathrm{F}} y_{\mathrm{I}-\mathrm{F}},$$

容易发现

$$v_{n-k} = (-1)^k \sum a_{i_1} \cdots a_{i_k},$$

右端的求和是取遍集合 $\{1, 2, \cdots, n\}$ 的所有子集 $\{i_1, \cdots, i_k\}$ 上的. 由于

$$u_{n-k} = u_n v_{n-k},$$

由 (11) 和 (12) 得到

$$\sum a_{i_1} \cdots a_{i_k} = (-1)^k u_{n-k}/u_n; \tag{13}$$

称此式为方程 $f(x) = 0$ 的**根与系数的关系**. (13) 式左端称为方程 $f(x) = 0$ 的根的**初等对称函数**. 参见本节的习题 13.

例 2 如果 u 和 v 是一个二次方程

$$ax^2+bx+c=0$$

的 (相异或相同的) 根, 则有

$$u+v=-b/a, \quad uv=c/a.$$

例 3 如果 u,v 和 w 是一个三次方程

$$ax^3+bx^2+cx+d=0$$

的 (相异或相同的) 根, 则有

$$u+v+w=-b/a, \quad uv+wu+uv=c/a, \quad uvw=-d/a.$$

我们注意到对于 $k=1$, 关系 (13) 写成

$$a_1+\cdots+a_n=-u_{n-1}/u_n, \tag{14}$$

对于 $k=n$, (13) 写成

$$a_1\cdots a_n=(-1)^n u_0/u_n. \tag{15}$$

§33 习题

¶1. 设 G_n 是满足方程 $z^n=1$ 的复数 z 的集合.

a) 证明 G_n 是非零复数乘法群 $\mathbf{C}^*$ 的一个 n 阶子群.

b) 设

$$z=\cos(2k\pi/n)+i\sin(2k\pi/n)$$

是 G_n 的一个元素. 证明 z 是 G_n 的一个生成元 (即单位的所有 n 次方根是 z 的一个幂) 必须并且只需 k 与 n 互素 (这时称 z 是单位的一个 n 次**原根**).

c) 不假定 k 与 n 互素, 证明 z 在群 G_n 内的阶 (即使得 $z^d=1$ 的最小的整数 $d\geqslant 1$) 是 $n/(k,n)$, (k,n) 表示 k 和 n 的最大公因子, 并且 z 是单位的 d 次原根.

d) 设 $\varphi(n)$ 是单位的 n 次原根的数目. 证明 $\varphi(n)$ 是满足条件 $1\leqslant k\leqslant n$ 的与 n 互素的整数的数目, 并且这也是环 $\mathbf{Z}/n\mathbf{Z}$ 的可逆元的数目. 证明

$$n=\sum_{d|n}\varphi(d),$$

求和取遍 n 的所有因子 d 上 (符号 $d|n$ 表示 d 整除 n). 证明

$$\varphi(n)=n\left(1-\frac{1}{p_1}\right)\cdots\left(1-\frac{1}{p_r}\right),$$

其中 $p_1,\cdots,p_r$ 是 n 的不同的素因子.

e) 对于 $n = 2, 3, 4, 6, 8, 12, 16, 20, 24$, 按照单位的 n 次根的阶分类. 计算单位的 24 次根的实部和虚部.

¶2. 设 K 是一个交换域, 而 n 是一个整数, 使得 $x^n = 1$ 在 K 内有 n 个根. 证明由这个方程的根组成的乘法群 K^* 的 n 阶子群是*循环群* (利用 §7 的习题 20).

由此推出, 如果 K 是有 q 个元素的有限域, 则乘法群 K^* 是循环群 (在 K 内考虑方程 $x^{q-1} = 1$). 作为特殊情形, 对于所有素数 p, 乘法群 $(\mathbf{Z}/p\mathbf{Z})^*$ 是循环群.

¶3. 设 K 是一个有 q 个元素的有限交换域, 而设 $a_1, \cdots, a_{q-1}$ 是 K 的非零元素. 证明如果用 X 表示 K 上的一个未定元, 则有

$$(X - a_1) \cdots (X - a_{q-1}) = X^{q-1} - 1$$

(利用 §33 的定理 1). 由此推出

$$a_1 \cdots a_{q-1} = -1$$

(为了确定符号, 会用到 q 是 K 的特征 p 的一个幂这个事实: §30, 习题 8).

取 $K = \mathbf{Z}/p\mathbf{Z}$, 由此推出 Wilson *定理*, 即对于素数 p 有

$$(p-1)! \equiv -1 (\bmod p).$$

¶4. 设 p 是一个素数, 而 r 是一个正整数. 称与 p 互素的一个整数 n 是一个模 p 的 **r 次剩余** (如果 $r = 2$, 则称为模 p **平方剩余**), 如果存在一个整数 x, 使得

$$x^r \equiv n \,(\bmod p).$$

利用乘法群 $(\mathbf{Z}/p\mathbf{Z})^*$ 是循环群这个事实 (习题 2) 证明: 如果 r 与 $p-1$ 互素, 则所有与 p 互素的 x 是一个模 p 的 r 次剩余. 如果 r 整除 $p-1$ (例如 $r = 2$, 而 p 是一个奇数, 这是最重要的情形), 一个与 p 互素的整数 n 是模 p 的一个 r 次剩余, 必须并且只需

$$n^{\frac{p-1}{r}} \equiv 1 \,(\bmod p).$$

模 p 的 r 次剩余的模 p 的剩余类的数目等于 $\dfrac{p-1}{r}$ (例如, 如果 p 是奇数, 则有 $\dfrac{p-1}{2}$ 个模 p 的平方剩余).

取 $p = 31$. 对于 $p - 1 = 30$ 的每一个因子 r 求模 p 的 r 次剩余.

¶5. (这个习题基于习题 1) 对于所有整数 $n \geqslant 1$, 称多项式

$$\Phi_n(X) = (X - \xi_1) \cdots (X - \xi_h)$$

为指标为 n 的**分圆多项式**, 其根是在域 $\mathbf{C}$ 内的单位的 $h = \phi(n)$ 个 n 次*原根*. 这个多项式一眼看来系数在 $\mathbf{C}$ 内, 但其实将证明其系数是*有理整数*. 我们约定 $\Phi_1(X) = X - 1$.

a) 如果 p 是素数, 则有

$$\Phi_p(X) = 1 + X + X^2 + \cdots + X^{p-1}.$$

b) 验证

$$\Phi_{12}(X) = X^4 - X^2 + 1.$$

c) 对于所有整数 $n \geqslant 1$ 有

$$X^n - 1 = \prod_{d|n} \Phi_d(X),$$

其中出现在右端的乘积是取遍 n 的所有因子 (含有 1 和 n) (利用多项式 $X^n - 1$ 的线性因式的乘积的分解).

d) 利用关系

$$\Phi_n(X) = \frac{X^n - 1}{\prod\limits_{\substack{d|n \\ d \neq n}} \Phi_d(X)}$$

并且进行关于 n 的归纳推理, 证明 Φ_n 的系数是有理整数. (这个结果推广到所有代数闭域 K, 只要限于不被 K 的特征整除的整数 n, 由于 §30 习题 14, 上述限制其实不是什么限制.)

¶¶6. 证明在整数 $n \geqslant 1$ 的集合上, 存在唯一的一个取整数值的函数 μ (Möbius **函数**), 满足关系

$$\sum_{d|n} \mu(d) = \begin{cases} 1, & \text{如果 } n = 1, \\ 0, & \text{如果 } n > 1 \end{cases}$$

(求和是取遍满足条件 $1 \leqslant d \leqslant n$ 的 n 的因子).

证明有

$$\mu(1) = 1;$$
$$\mu(p) = -1, \text{ 如果 } p \text{ 是素数};$$
$$\mu(p^r) = 0, \text{ 如果 } p \text{ 是素数, 并且 } r \geqslant 2.$$

证明

$$\mu(mn) = \mu(m)\mu(n), \text{ 如果 } m \text{ 和 } n \text{ 互素}. \tag{$*$}$$

(注意到当 m 和 n 互素时, mn 的所有因子以唯一的一种方式写成 m 的一个因子和 n 的一个因子的乘积, 假定 $(*)$ 已经对于满足 $m'n' < mn$ 的一对数 m', n' 建立, 采用归纳推理.) 由上面的结果推出

$$\mu(n) = \begin{cases} 1, & \text{如果 } n = 1, \\ (-1)^r, & \text{如果 } n \text{ 被素数的平方整除}, \\ 0, & \text{如果 } n \text{ 不被素数的平方整除}. \end{cases}$$

对于 $1 \leqslant n \leqslant 100$ 计算 $\mu(n)$.

¶¶7. 设 f 是整数 $n \geqslant 1$ 的集合上的在加法群 A 上取值的一个函数. 令

$$g(n) = \sum_{d|n} f(d)$$

定义一个新的函数 g, 其中的求和是取遍满足条件 $1 \leqslant d \leqslant n$ 的 n 的因子 d. 证明反之有

$$f(n) = \sum_{d|n} g(d)\mu\left(\frac{n}{d}\right),$$

其中的 μ 是上一个习题中的 Möbius 函数 (仅用到定义 μ 的性质).

如果 A 是按乘法写出的交换群, 前面的公式如何修改?

¶¶8. (这个习题基于习题 5, 6 和 7) 证明分圆多项式由下列关系给定:

$$\Phi_n(\mathrm{X}) = \prod_{d|n} (\mathrm{X}^d - 1)^{\mu(\frac{n}{d})}.$$

对于 $n = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15$ 计算多项式 $\Phi_n(\mathrm{X})$, 并且证明

$$\begin{aligned}\Phi_n(\mathrm{X}) = {} & \mathrm{X}^{48} + \mathrm{X}^{47} + \mathrm{X}^{46} - \mathrm{X}^{43} - \mathrm{X}^{42} - 2\mathrm{X}^{41} - \mathrm{X}^{40} - \mathrm{X}^{39} + \mathrm{X}^{36} + \mathrm{X}^{35} + \mathrm{X}^{34} + \mathrm{X}^{33} \\ & + \mathrm{X}^{32} + \mathrm{X}^{31} - \mathrm{X}^{28} - \mathrm{X}^{26} - \mathrm{X}^{24} - \mathrm{X}^{22} - \mathrm{X}^{20} + \mathrm{X}^{17} + \mathrm{X}^{16} + \mathrm{X}^{15} + \mathrm{X}^{14} + \mathrm{X}^{13} \\ & + \mathrm{X}^{12} - \mathrm{X}^{9} - \mathrm{X}^{8} - 2\mathrm{X}^{7} - \mathrm{X}^{6} - \mathrm{X}^{5} + \mathrm{X}^{2} + \mathrm{X} + 1,\end{aligned}$$

假定 Faddeev 和 Sominskii 是正确的, 作者没有验证(*).

9. 分解有理分式

$$\frac{f(\mathrm{X})}{\mathrm{X}^n - 1}$$

为复部分分式的和, 其中 f 是复系数的任意一个多项式.

¶10. 设 $z_1, \cdots, z_n$ 是在 **C** 内的单位的 n 次根. 证明

$$z_1^h + \cdots + z_n^h = \begin{cases} n, & \text{如果 } h \equiv 0 (\mathrm{mod}\ n), \\ 0, & \text{如果 } h \not\equiv 0 (\mathrm{mod}\ n) \end{cases}$$

(左端乘以 z_1^h).

11. 用 $z_1, \cdots, z_n$ 表示在 **C** 内的单位的 n 次根. 证明下列关系:

$$\prod_{k=1}^{n} (a + bz_k) = a^n + (-1)^{n-1} b^n,$$

$$\prod_{k=1}^{n} (z_k^2 - 2z_k \cos\theta + 1) = 2(1 - \cos n\theta),$$

$$\prod_{k=1}^{n} \frac{(t + z_k)^n - 1}{t} = \prod_{k=1}^{k=n-1} [t^n - (z_k - 1)^n].$$

12. 设 u, v, w 是复系数的三次方程

$$ax^3 + bx^2 + cx + d = 0$$

的三个 (不同或相等的) 根. 利用 §33 第 6 小节例 4 借助 a, b, c, d 计算表达式

$$u^2 + v^2 + w^2, \quad u^3 + v^3 + w^3.$$

(*) Maple 给出相等的结果 > cyclotomic(105, x);

$$\begin{aligned}& 1 + x + x^2 - x^5 - x^8 - 2x^7 + x^{35} - x^{28} + x^{48} + x^{46} - x^{43} - 2x^{41} - x^{40} - x^{39} + x^{36} + x^{34} + x^{33} \\ & + x^{31} - x^{26} - x^{24} - x^{22} - x^{20} + x^{17} + x^{16} + x^{15} + x^{14} + x^{12} - x^{9} - x^{6} + x^{47} - x^{42} + x^{32} + x^{13}\end{aligned}$$

—— 译者注

¶13. 设 $X_1, \cdots, X_n$ 是交换环 K 上的未定元. 称多项式

$$\begin{aligned}
s_1 &= X_1 + \cdots + X_n = \sum X_i, \\
s_2 &= X_1X_2 + \cdots + X_{n-1}X_n = \sum_{1\leqslant i<j\leqslant n} X_iX_j, \\
s_3 &= X_1X_2X_3 + \cdots = \sum_{1\leqslant i<j<k\leqslant n} X_iX_jX_k, \\
&\cdots\cdots\cdots\cdots \\
s_n &= X_1X_2\cdots X_n
\end{aligned}$$

为 $X_1, \cdots, X_n$ 的**初等对称函数** (这些表达式涉及用代数方程的根计算其系数, 参见 §33 的第 6 小节). 另外, 称一个多项式 $f \in K[X_1, \cdots, X_n]$ 是**对称的**, 如果对于整数 $1, \cdots, n$ 的所有置换 s 有

$$f(X_{s(1)}, \cdots, X_{s(n)}) = f(X_1, \cdots, X_n).$$

a) 证明所有的对称多项式是系数在 K 内的 $s_1, \cdots, s_n$ 的一个多项式. [可以关于整数 $n + d^\circ(f)$ 进行归纳推理. 首先注意 $f(X_1, \cdots, X_{n-1}, 0)$ 关于 $X_1, \cdots, X_{n-1}$ 是对称的, 从而是 $X_1, \cdots, X_{n-1}$ 的初等对称函数的一个多项式, 这些初等对称函数由在 $X_1, \cdots, X_n$ 的初等对称函数中令 $X_n = 0$ 而得到. 由此推出存在一个其次数至多等于 f 的次数的多项式 $p(s_1, \cdots, s_{n-1})$, 使得

$$g(X_1, \cdots, X_n) = f(X_1, \cdots, X_n) - p(s_1, \cdots, s_{n-1})$$

当 $X_n = 0$ 时等于 0. 考虑到 g 的对称性由此推出

$$g(X_1, \cdots, X_n) = X_1 \cdots X_n h(X_1, \cdots, X_n),$$

其中的 h 是对称的, 并且 $d^\circ(h) < d^\circ(f)$.]

b) 证明: 如果 $p \in K[X_1, \cdots, X_n]$ 满足 $p(s_1, \cdots, s_n) = 0$, 则 $p = 0$ (关于 n 进行归纳推理. 取 p 为 s_n 的最小可能的次数, 在结果中令 $X_n = 0$).

c) 由此推出对称多项式 f 借助 $s_1, \cdots, s_n$ 的表达式是唯一的.

d) 假定 f 关于 $X_1, \cdots, X_n$ 是总次数 k 次齐次的, 令

$$f(X_1, \cdots, X_n) = p(s_1, \cdots, s_n),$$

证明实际出现在 p 内的单项式仅是这样的

$$s_1^{r_1} \cdots s_n^{r_n},$$

其中的指数满足关系

$$r_1 + 2r_2 + \cdots + nr_n = k.$$

e) 借助初等对称函数计算下列多项式:

$$\mathrm{X}_1^2\mathrm{X}_2+\mathrm{X}_1\mathrm{X}_2^2+\mathrm{X}_1^2\mathrm{X}_3+\mathrm{X}_1\mathrm{X}_3^2+\mathrm{X}_2^2\mathrm{X}_3+\mathrm{X}_2\mathrm{X}_3^2 \quad (n=3);$$
$$(2\mathrm{X}_1-\mathrm{X}_2-\mathrm{X}_3)(2\mathrm{X}_2-\mathrm{X}_1-\mathrm{X}_3)(2\mathrm{X}_3-\mathrm{X}_1-\mathrm{X}_2) \quad (n=3);$$
$$(\mathrm{X}_1\mathrm{X}_2+\mathrm{X}_3\mathrm{X}_4)(\mathrm{X}_1\mathrm{X}_3+\mathrm{X}_2\mathrm{X}_4)(\mathrm{X}_1\mathrm{X}_4+\mathrm{X}_2\mathrm{X}_3) \quad (n=4);$$
$$\sum_{i\neq j}\mathrm{X}_i^3\mathrm{X}_j^2;\quad \sum_{i\neq j\neq k}\mathrm{X}_i^3\mathrm{X}_j^2\mathrm{X}_k;\quad \sum_{i\neq j\neq k}\mathrm{X}_i^4\mathrm{X}_j\mathrm{X}_k;\quad \sum_{s\in\mathfrak{S}_n}(a_1\mathrm{X}_{s(1)}+\cdots+a_n\mathrm{X}_{s(n)});$$
$$\sum_{\substack{i>k\\ j\neq i,j\neq k}}(\mathrm{X}_i+\mathrm{X}_k-\mathrm{X}_j)^2\ (n\ 任意).$$

¶14. 沿用上题的记号, 考虑 **Newton 和**

$$\sigma_k(\mathrm{X}_1,\cdots,\mathrm{X}_n)=\mathrm{X}_1^k+\cdots+\mathrm{X}_n^k \quad (k=0,1,\cdots).$$

证明可以借助下列公式用 $s_1,\cdots,s_n$ 表示它们:

$$\sigma_k-s_1\sigma_{k-1}+s_2\sigma_{k-2}-\cdots+(-1)^{k-1}s_{k-1}\sigma_1+(-1)^k s_{k-1}\sigma_0=0 \quad 对于\ k<n,$$
$$\sigma_k-s_1\sigma_{k-1}+\cdots+(-1)^n s_n\sigma_{k-n}=0 \quad 对于\ k\geqslant n.$$

借助初等对称函数 s_i 对于 $0\leqslant k\leqslant 6$ 计算 Newton 和 σ_k.

¶15. 设

$$x^n+a_{n-1}x^{n-1}+\cdots+a_0=0$$

是一个系数在一个交换域 K 内的代数方程. 设 $x_1,\cdots,x_n$ 是它在 K 的代数闭扩张 (可以取它为 **C**, 而 K 是 **C** 的一个子域, 但是显然这个假设没有化简什么) 内的根. 设 f 是系数在 K 内的 n 个未定元的一个多项式. 证明存在一个系数在 K 内的 n 个未定元的一个多项式 p, 使得

$$f(x_1,\cdots,x_n)=p(a_{n-1},\cdots,a_0),$$

并且 p 仅依赖于 f (利用习题 13). 应用:

a) 计算方程

$$x^6-4x^5+3x^3-4x^2+x+1=0$$

的根的五次幂的和:

b) 计算和

$$\sum x_i^2x_j^2x_kx_h,$$

其中 $x_1,\cdots,x_5$ 表示以下方程的根:

$$x^5-4x^3+x^2+3x+1=0.$$

c) 用 x_1,x_2,x_3 表示

$$x^3+ax^2+bx+c=0$$

的根, 组成方程, 其根是下列各值:

i) x_1+x_2,x_2+x_3,x_3+x_1,

ii) $x_1^2 - x_2x_3, x_2^2 - x_3x_1, x_3^2 - x_1x_2$,

iii) $(x_1 + jx_2 + j^2x_3)^3, (x_1 + j^2x_2 + jx_3)^3$, 其中 $j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$.

¶16. 设

$$x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0 \qquad (*)$$

是一个系数在交换域 K 内的代数方程. 用 $x_1, \cdots, x_n$ 表示它的在 K 的一个代数闭扩张内的 (不同的或相同的) 根. 称表达式

$$\mathrm{D} = \prod_{1 \leqslant i < j \leqslant n} (x_i - x_j)^2$$

为给定方程的**判别式**.

a) 证明存在系数在 K 内的 n 个未定元并且不依赖方程 $(*)$ 的一个多项式 p, 使得

$$\mathrm{D} = p(a_{n-1}, \cdots, a_0).$$

b) 计算二、三、四次方程的判别式.

c) 方程 $(*)$ 至少具有一个二重根, 必须并且只需它的判别式为零.

d) 确定 λ 的值, 对于它们下列方程至少具有一个重根:

$$x^2 - 3x + \lambda = 0;$$
$$x^3 - 8x^2 + (13 - \lambda)x - 6 - 2\lambda = 0;$$
$$x^4 - 4x^3 + (2 - \lambda)x^2 + 2x - 2 = 0.$$

¶¶e) 证明方程

$$x^n + px + q = 0$$

的判别式等于

$$(-1)^{\frac{n(n-1)}{2}} n^n q^{n-1} + (-1)^{\frac{(n-1)(n-2)}{2}} (n-1)^{n-1} p^n.$$

f) 用 f 表示出现在方程 $(*)$ 左端的多项式. 证明方程 $(*)$ 的判别式 D 还由下面的公式给定:

$$(-1)^{\frac{n(n-1)}{2}} \mathrm{D} = \prod_{1 \leqslant i \leqslant n} f'(x_i).$$

g) 证明方程

$$\frac{x^n}{n!} + \frac{x^{n-1}}{(n-1)!} + \cdots + x + 1 = 0$$

的判别式等于

$$(-1)^{\frac{n(n-1)}{2}}.$$

¶¶h) 考虑 (习题 5) 分圆方程

$$\Phi_n(x) = 0.$$

证明它的判别式等于

$$(-1)^{\frac{\varphi(n)}{2}} \frac{n^{\varphi(n)}}{\prod_{p|n} p^{\frac{\varphi(n)}{p-1}}}$$

(利用问题 f) 和习题 8).

17. 设 K 是一个交换域, 而 f 是系数在 K 内的一个未定元的多项式. 我们打算证明: 存在 K 的一个扩张域 L (即 K 是 L 的一个子域), 使得在 L 内 f 至少具有一个根 (这个结果是 Steinitz 定理证明的第一步).

a) 证明可以归结为 f 是在 K 上不可约的这种情形.

b) 在多项式环 $\mathrm{A} = \mathrm{K}[\mathrm{X}]$ 内考虑由 f 生成的理想 $\mathrm{I} = (f)$, 并且组成商环 $\mathrm{L} = \mathrm{A}/\mathrm{I}$ (§8, 习题 7). 证明这是一个域.

c) 设 j 是从 K 到 L 内的映射, 它令每一个 $c \in \mathrm{K}$ 对应常多项式 c 模 I 的类. 证明这是从 K 到 L 的一个子域上的同构 (下面我们约定等同 $c \in \mathrm{K}$ 和它在 L 内的像 $j(c)$).

d) 设 $z \in \mathrm{L}$ 是在从 K[L] 到 L 上的典范映射下的像. 证明 z 是 f 的根, 并且 $\mathrm{L} = \mathrm{K}[z]$. (这就证明了开始所宣布的结果.)

e) 取 $f(\mathrm{X}) = \mathrm{X}^2 - d$, 其中 $d \in \mathrm{K}$ 不是在 K 内的一个平方. 证明上面的构造归结为 §9 中的构造. 作为特殊情形, 验证域 $\mathbf{C}$ 是环 $\mathbf{R}[\mathrm{X}]$ 对于由 $\mathrm{X}^2 + 1$ 生成的理想的商域 ($\mathbf{C}$ 的这个构造方法属于 Cauchy).

f) 取 $f(\mathrm{X}) = \mathrm{X}^3 + p\mathrm{X} + q$, 假定它在 K 上是不可约的. 对于对应的域 L 给出与对于 §9 中的环 $\mathrm{K}[\sqrt{d}]$ 所给出的描述相类似的描述.

g) 如果多项式 f 不是不可约的, 在前述的构造中会出现什么现象?

¶¶18. 设 $f_1, \cdots, f_n$ 是系数在一个交换域 K 内的一个未定元的非常元的多项式. 证明在环 $\mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$ 内存在一个含有 $f_1(\mathrm{X}_1), \cdots, f_n(\mathrm{X}_n)$ 的极大理想 (利用 §27 的习题 15). 像前一个习题那样推理, 从而推出存在 K 的这样的代数扩张 L, 在 L 内每一个 f_i 至少有一个根.

(Steinitz 定理的完全证明是这个习题的推理的直接推广: 引进一个有无穷多个变量的多项式环, 其变量跟首项系数为 1 的系数在 K 内的不可约多项式一样多, 然后取这个环关于所选择的一个极大理想的商.)

¶¶19. 保留上题的记号, 证明环 $\mathrm{K}[\mathrm{X}_1, \cdots, \mathrm{X}_n]$ 的含有 $f_1(\mathrm{X}_1), \cdots, f_n(\mathrm{X}_n)$ 的所有素理想是极大的 (参见 §26, 习题 3).

¶20. 设 K 是一个交换域. 称 K 的一个扩张 L (即以 K 为子域的一个交换域) 是**代数的**, 如果所有 $x \in \mathrm{L}$ 在 K 上是代数的. 证明 K 是代数闭的, 必须并且只需对于 K 的所有代数扩张 L 有 L=K.

21. 一个代数闭域总具有无穷多个元素.

¶22. (d'Alembert-Gauss 定理的证明) 这个习题假定知晓平面内的连续函数的性质 (尤其是事实: 一个在一个紧集上的正的连续函数取到最小值). 用

$$f(z) = a_0 + \cdots + a_n z^n$$

表示复系数的非常值的一个多项式, 假定 $a_n \neq 0$.

a) 证明比值

$$f(z)/a_n z^n$$

当 $|z|$ 无限增加时趋于 1, 即对于所有 $\varepsilon > 0$ 存在 $r > 0$, 使得

$$|z| > r \quad \text{蕴含} \quad \left|1 - \frac{f(z)}{a_n z^n}\right| < \varepsilon.$$

b) 设

$$m = \inf_{z \in \mathbf{C}} |f(z)|,$$

证明存在一个数 $r' > 0$, 使得

$$|z| > r' \quad 蕴含 \quad |f(z)| \geqslant m + 1.$$

应用本题开始提到的最小值定理到在紧集 $|z| \leqslant r'$ 上的连续函数 $|f(z)|$, 证明存在 $z_0 \in \mathbf{C}$, 使得

$$|f(z_0)| = m.$$

[如果 d'Alembert-Gauss 定理成立, 显然 $m = 0$. 为了证明所提到的定理成立, 必须并且显然只需证明 $m = 0$. 这正是下一个问题的目的.]

c) 假定 $m \neq 0$. 用 $z + z_0$ 代替 z, 用 $f/f(z_0)$ 代替 f, 就归结为情形:

$$f(0) = 1, \quad |f(z)| \geqslant 1 \quad 对于所有 \ z \in \mathbf{C}.$$

设

$$f(z) = 1 + b_q z^q + b_{q+1} z^{q+1} + \cdots + b_n z^n, \quad 其中 \ b_q \neq 0.$$

证明存在一个数 $\mathrm{M} > 0$, 使得

$$|z| \leqslant 1 \quad 蕴含 \quad |f(z) - 1 - b_q z^q| \leqslant \mathrm{M}|z|^{q+1}.$$

由此推出只要选择 z, 使它满足

$$|z| \leqslant 1, \quad |z| < |b_q|/\mathrm{M}, \quad \mathrm{Arg}(b_q) + q \cdot \mathrm{Arg}(z) = \pi,$$

就有 $|f(z)| < 1$ (这与假设矛盾, 从而完成了证明).

¶¶23. 设 E 是一个交换代数闭域, L 是任意一个交换域, 而 σ 是从 L 到 E 的一个子域 L$'$ 的一个同构. 考虑 L 的一个扩张 M, 并且假定 $\mathrm{M} = \mathrm{L}[z]$, 其中 z 在 L 上是代数的. 设

$$f(\mathrm{X}) = \mathrm{X}^n + a_{n-1}\mathrm{X}^{n-1} + \cdots + a_0$$

是 z 在 L 上的极小多项式, 把从它通过 σ 得到的系数在 L$'$ 内的多项式记作

$$f^\sigma(\mathrm{X}) = \mathrm{X}^n + \sigma(a_{n-1})\mathrm{X}^{n-1} + \cdots + \sigma(a_0),$$

考虑 f^σ 在 E 内的一个根 z', 设 $\mathrm{M}' = \mathrm{L}'[z']$. 证明存在唯一的一个从 M 到 M$'$ 上的同构 σ', 它在 L 上与 σ 重合, 并且把 z 映射到 z'.

由此推出下列结果: 设 E 是一个交换域 K 的代数闭扩张, 而 L 是 K 的有限扩张. 则存在从 L 到 E 的一个子域上的同构 j, 使得对于所有 $x \in \mathrm{K}$ 有 $j(x) = x$. (关于 L 在 K 上的次数进行归纳推理. 事实上可以证明这个结果对于所有代数扩张都成立, 不论扩张是有限的还是无限的.)

¶24. 设 K 是一个交换域, E 是 K 的一个代数闭扩张, 而 L 是 K 的一个有限扩张. 假定在 K 上 L 是可分的 (§26, 习题 4, h)), 并且令 $n = [\mathrm{L} : \mathrm{K}]$. 证明从 L 到 E 内的对于所有 $x \in \mathrm{K}$ 有 $j(x) = x$ 的同构的数目刚好是 n (像在习题 26 中那样推理, 并且利用 §32 的习题 11).

用 $j_1, \cdots, j_n$ 表示所提到的 n 个同构. 对于 $k \neq h$, 用 L_{hk} 表示使得 $j_k(z) = j_h(z)$ 成立的 $z \in \mathrm{L}$ 的集合. 证明这是包含 K 的并且异于 L 的 L 的一个子域.

假定 K 是无限的, 证明 L_{kh} 的并集不是整个 L, 并且存在一个 $z \in \mathrm{L}$, 使得 n 个元素 $j_1(z), \cdots, j_n(z)$ 两两不同. 由此推出, 如果 L 是无限域 K 的一个有限可分的扩张, 则存在 $z \in \mathrm{L}$, 使得 $\mathrm{L} = \mathrm{K}[z]$ (**本原元素定理**, 首先由 Dedekind 对于代数数域即 $\mathrm{K} = \mathbf{Q}$ 的情形证明. 事实上它对于有限的 K 也有效, 参见上面的习题 2).

¶¶25. 考虑 $\mathbf{Q}$ 的扩张

$$\mathrm{L} = \mathbf{Q}[\sqrt{3}, \sqrt[3]{2}].$$

构造一个代数数 z, 使得

$$\mathrm{L} = \mathbf{Q}[z].$$

¶¶26. 设 K 是一个交换域, L 是 K 的 n 次可分扩张, E 是 K 的代数闭扩张, 而 $j_1, \cdots, j_n$ 是从 L 到 E 内的同构, 对于所有 $x \in \mathrm{K}$ 有 $j_k(x) = x$ (习题 27). 我们要证明对于所有 $z \in \mathrm{L}$ 有

$$\mathrm{Tr}_{\mathrm{L}/\mathrm{K}}(z) = j_1(z) + \cdots + j_n(z),$$
$$\mathrm{N}_{\mathrm{L}/\mathrm{K}}(z) = j_1(z) \cdots j_n(z).$$

a) 设 $(a_i)_{1 \leqslant i \leqslant n}$ 是在 K 上 L 的基. 证明 (元素在 E 内的) 矩阵

$$\mathrm{A} = (j_k(a_h))_{1 \leqslant h,k \leqslant n}$$

是可逆的 (利用 §§10, 11 的习题 16 以及 Cramer 方程组的特征).

b) 对于所有 $z \in \mathrm{L}$ 令

$$za_i = \sum_j \xi_{ij} a_j,$$

其中的 $\xi_{ij} \in \mathrm{K}$. 引进矩阵

$$\mathrm{U}_z = (\xi_{ij})_{1 \leqslant i,j \leqslant n}$$

和

$$\mathrm{D}_z = \begin{pmatrix} j_1(z) & 0 & \cdots & 0 \\ 0 & j_2(z) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & j_n(z) \end{pmatrix},$$

证明有关系

$$\mathrm{U}_z = \mathrm{A}\mathrm{D}_z\mathrm{A}^{-1}.$$

c) 注意到 $\mathrm{Tr}_{\mathrm{L}/\mathrm{K}} = \mathrm{Tr}(\mathrm{U}_z)$ 和 $\mathrm{N}_{\mathrm{L}/\mathrm{K}} = \det(\mathrm{U}_z)$ 完成证明 (参见 §26, 习题 4).

[读者会注意到这个推理还证明了 L 的自同态

$$u_z : x \to zx$$

的特征值正是 $j_k(z), 1 \leqslant k \leqslant n$, 这里把 L 看作 K 上的 n 维向量空间.]

27. 证明多项式

$$\mathrm{X}^{n_1} + \mathrm{X}^{n_2} + \cdots + \mathrm{X}^{n_k}$$

是被多项式

$$1+\mathrm{X}+\mathrm{X}^2+\cdots+\mathrm{X}^{k-1}$$

整除的, 其中 $n_i \equiv r-1(\bmod k)$.

¶28. 对于 r 的哪些值多项式 $\Phi_n(\mathrm{X}^r)$ 被 $\Phi_n(\mathrm{X})$ 整除?

29. 设 f 是系数在一个交换域内的一个未定元的多项式. 如果 $f(\mathrm{X}^n)$ 被 $\mathrm{X}-1$ 整除, 则 $f(\mathrm{X}^n)$ 被 X^n-1 整除.

¶¶30. (Hilbert 零点定理的证明)

a) 设 L 是一个交换域而 A 是 L 的一个子环. 假定存在 L 的有限个元素 $y_1,\cdots,y_q$, 使得

$$\mathrm{L}=\mathrm{A}[y_1,\cdots,y_q],$$

并且设每一个 y_j 满足一个系数在 A 内的非平凡的代数关系. 证明存在 A 的一个元素 $b\neq 0$, 使得

$$\mathrm{L}=\mathrm{A}[b^{-1}]$$

是一个交换域. (选择 b 使得 L 是一个有限生成的 $\mathrm{A}[b^{-1}]$-模, 并且应用 §19 的习题 24.)

b) 证明环 A 的所有非零素理想含有 b.

c) 假定存在 L 的一个子域 K, 使得 A 是由 K 和在 K 上代数无关的 L 的有限个元素生成的 L 的子环. 证明此时有 A=K, 并且 L 是 K 的有限代数扩张 (注意, 在一个域上的多项式环内非零素理想的交集缩减为 0).

d) 设 K 是一个交换域, 而 L 是 K 的一个扩张. 假定存在 L 的有限个元素 $z_1,\cdots,z_r$, 使得

$$\mathrm{L}=\mathrm{K}[z_1,\cdots,z_r].$$

证明此时 L 是 K 的有限的代数扩张, 并且作为特殊情形, 如果 K 是代数闭域, 则有 L=K [从集合 $\{z_1,\cdots,z_r\}$ 抽取尽可能多的代数无关的元素, 并且把 c) 应用到由 K 和这些元素生成的环 A].

e) 设 K 是一个交换域, 而 $\mathfrak{M}$ 是多项式环 $\mathrm{K}[\mathrm{X}_1,\cdots,\mathrm{X}_r]$ 的一个极大理想. 证明商环 $\mathrm{L}=\mathrm{K}[\mathrm{X}_1,\cdots,\mathrm{X}_r]/\mathfrak{M}$ 是 K 的有限扩张 [应用问题 d) 和 §8 的习题 7] . 由此推出零点定理 [参见 §34 习题 51 的另一个证明].

31. 设 p 是一个素数, 并且设 $\mathrm{K}=\mathbf{Z}/p\mathbf{Z}$ 是模 p 整数域. 如果

$$f(\mathrm{X})=f(\mathrm{X}_1,\cdots,\mathrm{X}_n)$$

是系数在 K 内的 n 个未定元的一个多项式, 用 $\mathrm{S}(f)$ 表示 f 的值的和, 即

$$\mathrm{S}(f)=\sum_{x_i\in\mathrm{K}} f(x_1,\cdots,x_n).$$

a) 假定 f 是一个单项式 $\mathrm{X}_1^{m_1}\cdots\mathrm{X}_n^{m_n}$. 证明 $\mathrm{S}(f)=0$, 除非所有 m_i 被 $p-1$ 整除并且 $\geqslant 1$, 在这种情形有 $\mathrm{S}(f)=(-1)^n$. (归结为一个未定元的情形.)

b) 利用 a) 证明, 如果 $d^\circ(f)<n(p-1)$, 则 $\mathrm{S}(f)=0$.

c) 设 $\varphi(\mathrm{X})=\varphi(\mathrm{X}_1,\cdots,\mathrm{X}_n)$ 是系数在 K 内的一个多项式. 令

$$f(\mathrm{X})=1-\varphi(\mathrm{X})^{p-1},$$

证明我们有

$$f(x) = 1, \quad 如果\ \varphi(x) = 0, x \in \mathrm{K}^n,$$
$$f(x) = 0, \quad 如果\ \varphi(x) \neq 0, x \in \mathrm{K}^n.$$

由此推出 φ 在 K^n 内零点的数目 $\mathrm{N}(\varphi)$ 满足同余式

$$\mathrm{N}(\varphi) \equiv \mathrm{S}(f) \pmod p.$$

d) 假定 $d^\circ(\varphi) < n$. 由 b) 和 c) 推出 $\mathrm{N}(\varphi) \equiv 0 (\mathrm{mod}\ p)$. 作为特殊情形, 如果 φ 没有常项, 则 φ 至少有一个异于 $(0, \cdots, 0)$ 的零点 (Chevalley 定理).

e) 把上面的结果推广到有限个方程 $\varphi_\alpha(x) = 0$ 的情形, 其中 $\sum d^\circ(\varphi_\alpha) < n$. (取 $1 - \varphi_\alpha^{p-1}$ 的乘积作为 f.)

第七章 矩阵的化简

给定有限维向量空间 E 的一个自同态 u, 找 E 的一个基, 使得 u 关于这个基的矩阵尽可能简单十分必要. 可以期望的最 "简单" 的是对角矩阵. 解决这个问题的主要工具是在 §34 陈述的特征向量理论, 在许多情形这已经满足需要. 对于完全一般的自同态 u, 应当利用 §35 的十分复杂的推理, 初学者可以忽略它, 不过它们跟 §34 更为初等的结果一样在应用中 (比如在常系数的线性常微分方程中) 是不可或缺的.

§36 的目的是提供一类矩阵, 可以预先断定它们可以化成对角形. 这里的主要工具是在向量空间 E 上的 "标量积", 它类似于通常空间的经典的标量积, 并且可给出 "正交" 向量概念的意义. §36 的考虑还是 "二次曲面" (二次代数方程定义的曲面) 分类的基础, 不过我们在课文中不予讨论.

§34 特征值

1. 特征向量和特征值的定义

设 E 是交换域 K 上的一个向量空间, 而 u 是 E 的一个自同态. 称所有使得 $u(x)$ 正比于 x 的非零向量 $x \in \mathrm{E}$ 为 u 的**特征向量**. 显然由特征向量 x 生成的 E 的向量子空间 D (即 x 的倍向量的集合) 满足 $u(\mathrm{D}) \subset \mathrm{D}$; 反之, 如果 E 的过原点的直线 D 满足 $u(\mathrm{D}) \subset \mathrm{D}$, 则在 D 上的所有非零向量是 u 的一个特征向量.

称一个标量 $\lambda \in \mathrm{K}$ 是 u 的一个**特征值**, 如果存在一个非零向量 $x \in \mathrm{E}$, 使得

$$u(x) = \lambda x, \tag{1}$$

这时 x 是 u 的一个特征向量; 称它是**相伴于**特征值 λ 的一个特征向量.

请注意关系 (1) 说明 x 被同态

$$u - \lambda \cdot 1$$

(这里 1 表示 E 的恒等同态) 变为零. 故 λ 是 u 的特征值, 必须并且只需

$$\operatorname{Ker}(u - \lambda \cdot 1) \neq \{0\}.$$

假定 E 在 K 上是有限维的, 就可以对于 $u - \lambda \cdot 1$ 应用 §23 的定理 8 的推论 2, 从而得到下列结果:

定理 1 设 u 是交换域 K 上的有限维向量空间 E 的一个自同态. 一个标量 $\lambda \in \mathrm{K}$ 是 u 的一个特征值, 必须并且只需

$$\det(u - \lambda \cdot 1) = 0. \tag{2}$$

这个结果将让我们可以指出 u 的特征值是系数在 K 内的一个代数方程的根.

2. 矩阵的特征多项式

设 E 是域 K 上的 n 维向量空间, 而 u 是 E 的一个自同态. 选择 E 的一个基 $(a_i)_{1 \leqslant i \leqslant n}$, 并且设

$$\mathrm{U} = (\alpha_{ij})_{1 \leqslant i,j \leqslant n}$$

是 u 关于这个基的矩阵 (§12, 第 3 小节). 由于 E 的恒等同态的矩阵是单位矩阵

$$1_n = (\delta_{ij})_{1 \leqslant i,j \leqslant n}, \text{ 这里 } \delta_{ij} = \begin{cases} 0, & \text{如果 } i \neq j, \\ 1, & \text{如果 } i = j. \end{cases}$$

我们发现自同态 $u - \lambda \cdot 1$ 关于所考虑的基由以下矩阵表示:

$$\mathrm{U} - \lambda \cdot 1_n = (\alpha_{ij} - \lambda \delta_{ij}) = \begin{pmatrix} \alpha_{11} - \lambda & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22-\lambda} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} - \lambda \end{pmatrix}, \tag{3}$$

它从 U 的对角元素 α_{ii} 减去 λ 而得到. 根据定理 1, u 的特征值满足关系

$$\begin{vmatrix} \alpha_{11} - \lambda & \alpha_{12} & \cdots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} - \lambda \end{vmatrix} = 0. \tag{4}$$

我们把 K 嵌入到系数在 K 内的一个未定元的多项式环 K[X] 内, 并且组成元素在交换环 K[X] 内的矩阵 $U - X \cdot 1_n$, 它的行列式

$$p_U(X) = \det(U - X \cdot 1_n) = \begin{vmatrix} \alpha_{11} - X & \alpha_{12} & \cdots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} - X \end{vmatrix} = 0 \tag{5}$$

是 K[X] 的一个元素, 即是系数在 K 内的一个未定元的多项式, 称它为**矩阵 U 的特征多项式**, 并且关系 (4) 表明 u 的特征值是方程

$$p_U(\lambda) = 0 \tag{6}$$

的根.

必须注意, 如果用 u 关于另一个基的矩阵 U' 代替 U 多项式 p_U 不变. 事实上, 我们对于一个矩阵 $P \in GL(n, K)$ 有 (§15, 定理 2 的推论)

$$U' = PUP^{-1},$$

那么由于 $P1_nP^{-1} = 1_n$, 我们有

$$U' - X \cdot 1_n = PUP^{-1} - XP1_nP^{-1} = P(U - X \cdot 1_n)P^{-1},$$

因此根据行列式乘积的定理有

$$p_{U'}(X) = \det(P) \cdot p_U(X) \cdot \det(P^{-1}) = p_U(X);$$

这就证明了我们的断言.

于是称多项式 $p_U(X)$ 为**同态 u 的特征多项式**, 其中 U 是关于 E 的任意一个基的矩阵. 将这个多项式记为

$$p_u(X),$$

显然有

$$p_u(\lambda) = \det(u - \lambda \cdot 1), \quad 对于所有\ \lambda \in K \tag{7}$$

(如果 K 是无限域, 那么根据 §28 的定理 1 这个关系足以刻画 p_u). 定理 1 表明 u 的特征值是它的特征多项式的根.

另外, 前面的考虑使得下列定义是自然的: 称 K (或更一般的, K 的一个扩张) 的所有满足方程 (6) 的元素为系数在 K 内的**方阵 $U = (\alpha_{ij})_{1 \leqslant i,j \leqslant n}$ 的特征值**.

3. 特征多项式的形式

保持前面的记号, 我们打算获得关于多项式

$$p_{\mathrm{U}}(\mathrm{X}) = \begin{vmatrix} \alpha_{11}-\mathrm{X} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22}-\mathrm{X} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn}-\mathrm{X} \end{vmatrix}$$

的形式的某些信息. 当展开这个行列式时, 我们发现 $n!$ 项的和, "主项" 是所考虑的行列式的主对角线元素的乘积

$$(\alpha_{11}-\mathrm{X})(\alpha_{22}-\mathrm{X})\cdots(\alpha_{nn}-\mathrm{X}). \tag{8}$$

其余的项中的每一个也是行列式的 n 个元素的乘积, 其中至多有 $n-2$ 个因子在主对角线上. 因此, $p_{\mathrm{U}}(\mathrm{X})$ 是 (8) 和 X 的至多 $n-2$ 次的一个多项式的和. 由此得到 $p_{\mathrm{U}}(\mathrm{X})$ 的次数 $>n-2$ 的单项式仅是出现在 (8) 中的这样的单项式, 于是有关系

$$(-1)^n p_{\mathrm{U}}(\mathrm{X}) = \mathrm{X}^n - (\alpha_{11}+\alpha_{22}+\cdots+\alpha_{nn})\mathrm{X}^{n-1}+\cdots$$

没有写出的项的次数至多是 $n-2$.

故可以写出

$$(-1)^n p_{\mathrm{U}}(\mathrm{X}) = \mathrm{X}^n - \tau_1(\mathrm{U})\mathrm{X}^{n-1} + \tau_2(\mathrm{U})\mathrm{X}^{n-2} - \cdots + (-1)^n\tau_n(\mathrm{U}); \tag{9}$$

系数 $\tau_i(\mathrm{U}) \in \mathrm{K}$ 显然是矩阵 U 的元素 α_{ij} 的带有理整数系数的多项式函数. 我们已经看到

$$\tau_1(\mathrm{U}) = \alpha_{11}+\alpha_{22}+\cdots+\alpha_{nn}$$

是矩阵 U 的对角线元素的和, 称这个标量为矩阵 U 的**迹**(*), 经常用记号

$$\mathrm{Tr}(\mathrm{U})$$

表示它.

此外, 在 (9) 中令 $\mathrm{X}=0$, 并且考虑到 $p_{\mathrm{U}}(0)=\det(\mathrm{U})$ 这一显然的事实, 我们发现

$$\tau_n(\mathrm{U}) = \det(\mathrm{U}).$$

例 1 如果 $n=2$ 和

$$\mathrm{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

(*) 参见 §12, 习题 8; §16, 习题 5; §26, 习题 4 和 5; §34, 习题 18 和 27.

则有

$$p_{\mathrm{U}}(\mathrm{X}) = \begin{vmatrix} a-\mathrm{X} & b \\ c & d-\mathrm{X} \end{vmatrix} = (a-\mathrm{X})(d-\mathrm{X}) - bc = \mathrm{X}^2 - \mathrm{Tr}(\mathrm{U}) + \det(\mathrm{U}).$$

如果 $n=3$ 和

$$\mathrm{U} = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix},$$

则有

$$p_{\mathrm{U}}(\mathrm{X}) = \begin{vmatrix} a-\mathrm{X} & b & c \\ a' & b'-\mathrm{X} & c' \\ a'' & b'' & c''-\mathrm{X} \end{vmatrix} = -\mathrm{X}^3 + \mathrm{Tr}(\mathrm{U})\mathrm{X}^2 - \tau_2(\mathrm{U})\mathrm{X} + \det(\mathrm{U}),$$

留给读者计算 $\tau_2(\mathrm{U})$. 参见习题 40.

4. 特征值的存在性

给定系数在一个代数闭域 K 内的代数方程总在 K 内至少具有一个根 (根据代数闭域的定义本身), 显然有下列结果:

定理 2　*在代数闭域上的有限维非零向量空间的所有自同态至少具有一个特征值.*

显然同样其元素在一个代数闭域内的方阵至少具有一个特征值. 在初等的应用中, 这些结果特别适用于 $\mathrm{K} = \mathbf{C}$.



注 1　系数在非代数闭域 K 内的方阵完全可以没有任何在 K 内的特征值. 例如取 $\mathrm{K} = \mathbf{R}$ 和矩阵

$$\mathrm{U} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

在直角坐标系里, 它表示绕原点的角为 θ 的旋转. 它的特征值是方程

$$\begin{vmatrix} \cos\theta-\lambda & -\sin\theta \\ \sin\theta & \cos\theta-\lambda \end{vmatrix} = (\cos\theta-\lambda)^2 + \sin^2\theta = 0$$

的解. 为了这个方程有实根, 必须并且只需 θ 是 π 的整倍数; 否则, 根是复数

$$\cos\theta \pm i\cdot\sin\theta,$$

它不是实数.

设 E 是交换域 K 上的有限维向量空间, 而 u 是 E 的一个自同态. 如果多项式 p_u 的所有根都在 K 内, 则说 u 的所有特征值在 K 内, 即 p_u 可以写成

$$(-1)^n p_u = (\mathrm{X} - \lambda_1)^{r_1} \cdots (\mathrm{X} - \lambda_q)^{r_q},$$

其中的 λ_i 是 p_u 在 K 内的不同的根, 而 r_i 是它们的重数. 同样说一个矩阵的所有特征值在 K 内, 如果它的特征多项式的所有的根在 K 内. 如果基础域是代数闭域则总是这种情形.

例 2 取 $\mathrm{K} = \mathbf{R}$ 和一个 2 阶矩阵, 我们有

$$p_u(\mathrm{X}) = \mathrm{X}^2 - \mathrm{Tr}(\mathrm{U})\mathrm{X} + \det(\mathrm{U}),$$

故 U 的所有特征值是实数当且仅当

$$\mathrm{Tr}(\mathrm{U})^2 - 4 \cdot \det(\mathrm{U}) \geqslant 0.$$

5. 化成三角矩阵

系数在一个环内的矩阵是**三角的**, 如果它有形式

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_{nn} \end{pmatrix},$$

即对角线下方的元素全是零. 另外, 域 K 上的有限维向量空间 E 的一个自同态 u 是**可三角化的**, 如果存在 E 的一个基, 使得 u 关于这个基的矩阵是三角矩阵, 即存在 E 的一个基 $x_1, \cdots, x_n$, 使得有下列形式的关系

$$\begin{cases} u(x_1) = x_1\alpha_{11}, \\ u(x_2) = x_1\alpha_{12} + x_2\alpha_{22}, \\ \cdots\cdots\cdots\cdots \\ u(x_n) = x_1\alpha_{1n} + x_2\alpha_{2n} + \cdots + x_n\alpha_{nn}. \end{cases} \tag{10}$$

如果这样, 那么显然 u 至少在 K 内具有一个特征值, 即 α_{11}, 并且有

$$p_u(\mathrm{X}) = \begin{vmatrix} \alpha_{11} - \mathrm{X} & \cdots & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} - \mathrm{X} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \alpha_{nn} - \mathrm{X} \end{vmatrix},$$

于是根据 §24 例 2 有

$$p_u(\mathrm{X}) = (\alpha_{11} - \mathrm{X})(\alpha_{22} - \mathrm{X}) \cdots (\alpha_{nn} - \mathrm{X}). \tag{11}$$

由于 α_{ii} 在 K 内, 我们看到 u 的所有特征值在 K 内. 这个条件对于 u 可三角化是必要的 (并且当 K 是代数闭时, 总是满足的), 也是充分的, 即我们有

定理 3 设 E 是交换域 K 上的有限维向量空间, 而 u 是 E 的一个自同态. u 是可三角化的, 必须并且只需 u 的所有特征值都在 K 内.

只需证明条件的充分性. 为了初学者方便, 我们首先对于 K 是代数闭域的情形证明, 一般情形处理起来多少有些困难.

我们应当指出, 如果 K 是代数闭的, 则 E 的所有自同态是可三角化的. 设 u 是这样的一个同态. 由于 K 是代数闭的, u 至少具有一个特征值, 即存在一个 $\alpha_{11} \in \mathrm{K}$ 和一个非零向量 $x_1 \in \mathrm{E}$, 使得

$$u(x_1) = \alpha_{11} x_1. \tag{12}$$

设 D 是 x_1 生成的 E 的 1 维子空间, 而 F 是 D 在 E 内的补空间, 使得 $\mathrm{E} = \mathrm{D} \oplus \mathrm{F}$ (直和). F 的存在性由 §19 定理 2 的推论 2 推出. 用 p 表示把每一个 $x \in \mathrm{E}$ 平行于 D (§17, 第 4 小节) 投影到 F 内得到的从 E 到 F 内的同态, 并且用

$$v(x) = p(u(x)) \quad \text{对于所有 } x \in \mathrm{F} \tag{13}$$

定义 F 的一个自同态. 由于 F 是 $n-1$ 维的, 我们可以用归纳推理, 假定定理已经对于 F 和 v 成立, 故可以构造 F 的一个基 $x_2, \cdots, x_n$, 使得有下列关系

$$\begin{cases} v(x_2) = x_2\alpha_{22}, \\ v(x_3) = x_2\alpha_{23} + x_3\alpha_{33}, \\ \cdots\cdots\cdots\cdots \\ v(x_n) = x_2\alpha_{2n} + x_3\alpha_{3n} + \cdots + x_n\alpha_{nn}. \end{cases} \tag{14}$$

而关系 (13) 指出对于所有 $x \in \mathrm{F}$, 向量 $u(x)$ 和向量 $v(x)$ 仅相差 x_1 的一个标量倍, 特别有下列形式的关系

$$\begin{cases} u(x_2) = x_1\alpha_{12} + v(x_2), \\ u(x_3) = x_1\alpha_{13} + v(x_3), \\ \qquad\cdots\cdots\cdots\cdots \\ u(x_n) = x_1\alpha_{1n} + v(x_n). \end{cases} \tag{15}$$

由于 $x_2, \cdots, x_n$ 组成 F 的一个基, 那么显然 $x_1, x_2, \cdots, x_n$ 组成 E 的一个基, 而关系 (12), (14) 和 (15) 表明 u 关于这个基的矩阵是三角的. 定理因此在 K 是代数闭域的情形被确立.

现在过渡到任意交换域 K 的情形. 自然受到前面证明的启发, 要采用关于 E 的维数的归纳推理. 首先选择 u 的一个特征值 $\alpha_{11} \in \mathrm{K}$ 和一个对应的特征向量 $x_1 \in \mathrm{E}$, 重新得到关系 (12). 像上面一样, 用 D 表示 x_1 生成的 E 中的直线, 而 F 是 D 在 E 内的 $n-1$ 维补空间, 设 v 是由 (13) 定义 F 的自同态. 一切显然结为确定 v 是可三角化的. 而如果采用归纳推理, 我们可以假定定理已经对于 $n-1=\dim(\mathrm{F})$ 证明, 为了指出 v 是可三角化的, 一切都归结为确认 v 的所有特征值都在 K 内.

为此选择 F 的任意一个基 $y_2, \cdots, y_n$, 并设 V 是 v 关于这个基的矩阵. 由于有如下形式的关系

$$u(y_i) = \alpha_{1i}x_1 + v(y_i) \quad (2 \leqslant i \leqslant n),$$

显然关于 E 的基 $x_1, y_2, \cdots, y_n$, 自同态 u 有矩阵

$$\mathrm{U} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \beta_{22} & \cdots & \beta_{2n} \\ 0 & \beta_{32} & \cdots & \beta_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & \beta_{n2} & \cdots & \beta_{nn} \end{pmatrix},$$

其中我们引入了记号 $\mathrm{V} = (\beta_{ij})_{2\leqslant i,j\leqslant n}$. U 的对角线元素减去 X, 并且借助 §24, 例 2 计算这样得到的矩阵的行列式, 我们发现

$$p_{\mathrm{U}}(\mathrm{X}) = (\alpha_{11} - \mathrm{X})p_{\mathrm{V}}(\mathrm{X}).$$

因此, 多项式 $p_{\mathrm{V}}(\mathrm{X}) = p_v(\mathrm{X})$ 整除 $p_{\mathrm{U}}(\mathrm{X}) = p_u(\mathrm{X})$, 而根据假设后者的根都在 K 内, 故 p_v 的所有根也都在 K 内 (§33, 第 1 小节的末尾), 即 v 的所有特征根必然都在 K 内, 证明到此结束.

推论 1 设 E 是代数闭域上的有限维向量空间, 则 E 的所有自同态都是可三角化的.

推论 2 设 U 是一个元素在一个代数闭域 K 内的阶 $n \geqslant 1$ 方阵, 则存在一个矩阵 $\mathrm{P} \in \mathrm{GL}(n, \mathrm{K})$, 使得矩阵

$$\mathrm{PUP}^{-1}$$

是三角矩阵.

为了证明推论, 只需对于由 U 定义的 K^n 的自同态应用推论 1, 并且考虑到 §15 定理 2 的推论.

推论 3 设 U 是一个元素在一个交换域 K 内的阶 $n \geqslant 1$ 方阵. 以下性质是等价的:

a) 存在一个矩阵 $\mathrm{P} \in \mathrm{GL}(n, \mathrm{K})$, 使得矩阵 PUP^{-1} 是三角矩阵.

b) 矩阵 U 的特征值都在 K 内.

考虑由 U 定义的 K^n 的自同态 u, 性质 a) 表明 u 是可三角化的. 由于 $p_u = p_U$, 推论 3 由定理立刻得到.

一个矩阵 $U \in M_n(K)$ 如果满足性质 a), 则称它是在 K **上可三角化的**. 当然, 总存在一个可逆矩阵 P, 使得 PUP^{-1} 是三角矩阵, 但是通常 P 的元素不是在 K 内的, 而是在包含 K 的一个代数闭域内 (例如, 如果 $K = \mathbf{R}$, 一般被迫取 P 为一个复矩阵) —— 为了确认这一事实, 只需应用推论 2. 说在 K 上 U 是可三角化的, 意味着可以假定 P 的系数在 K 内.

正如下节将要看到的, 定理 3 可以大大改进, 将指出存在 E 的一个基, 使得关于这个基, u 的矩阵含有比三角形矩阵更多的零. 但这个一般结果的完整的证明比定理 3 更困难, 而定理 3 在许多应用中已经足够了.

6. 特征值都是单特征值的情形

定理 3 没有提到涉及 u 的特征值的重数的任何条件. 在实际中经常碰到一种情形, 不仅仅多项式 p_u 的所有的根都在 K 内, 而且它们都是单根. 这时有一个比定理 3 精确得多的结果.

在建立这个结果之前, 先要证明以下结果:

定理 4　设 u 是一个交换域上的向量空间的一个自同态, 而 $x_1, \cdots, x_n \in E$ 是 u 的相伴于两两不等的特征值 $\lambda_1, \cdots, \lambda_n$ 的特征向量. 则向量 $x_1, \cdots, x_n$ 是线性无关的.

设 F 是由 $x_1, \cdots, x_n$ 生成的 E 的向量子空间, 则关系

$$u(x_i) = \lambda_i x_i \quad (1 \leqslant i \leqslant n) \tag{16}$$

指出如果

$$x = \sum \xi_i x_i$$

是 F 的一个向量, 则向量

$$u(x) = \sum \xi_i u(x_i) = \sum \lambda_i \xi_i x_i$$

还在 F 内. 因此 F 在 u 下是稳定的, 并且可以考虑 u 限制到 F 上得到的自同态 v. 我们有 $v(x_i) = \lambda_i x_i$, 因此 λ_i 是 v 的特征值, 假定 λ_i 是两两不同的, 我们看到 v 至少有 n 个特征值. 但它们是一个次数等于 $\dim(F)$ 的方程的根, 故有

$$n \leqslant \dim(F).$$

由于 $x_1, \cdots, x_n$ 生成 F, 故得 (§19, 定理 10) 这些向量组成 F 的一个基, 从而必定是线性无关的, 这就完成了证明. (还可以参见本节的习题 38 和 39.)

注 2 特征值 λ_i 两两不同的假设对于保证定理 4 的成立是本质的. 例如, 取 u 是零映射或恒等映射, 那么设 $x_1, \cdots, x_n$ 是非零向量, 显然 $x_1, \cdots, x_n$ 是 u 的特征向量, 但由此断定任意向量空间的任意 n 个非零向量总是线性无关的是无稽之谈!



现在我们可以陈述本小节开头所宣布的结果:

定理 5 设 u 是一个交换域 K 上的有限维 $n \geqslant 1$ 的向量空间 E 的一个自同态. 假定 u 的特征多项式 p_u 在 K 内具有 n 个单根 $\lambda_1, \cdots, \lambda_n$. 那么存在 E 的一个基, 使得关于这个基 u 的矩阵是

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

对于每个 i, 选择 u 的一个相伴于特征值 λ_i 的特征向量 x_i. 根据前面的定理 4, 这 n 个向量是线性无关的, 由于它们的个数 $n = \dim(\mathrm{E})$, 它们组成 E 的一个基. 显然 u 关于这个基的矩阵正是定理所断言的形式.

推论 设 U 是元素在一个交换域 K 内的一个 n 阶方阵. 假定它的特征多项式 p_{U} 在 K 内具有 n 个单根 $\lambda_1, \cdots, \lambda_n$. 则存在一个矩阵 $\mathrm{P} \in \mathrm{GL}(n, \mathrm{K})$, 使得矩阵

$$\mathrm{PUP}^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

为了证明推论只需把定理 5 应用到以 U 为矩阵的 K^n 的自同态.

例 3 取 $n = 2$ 和 $\mathrm{K} = \mathbf{C}$, U 的特征方程是

$$\mathrm{X}^2 - \mathrm{Tr}(\mathrm{U})\mathrm{X} + \det(\mathrm{U}) = 0,$$

当且仅当

$$\mathrm{Tr}(\mathrm{U}) - 4 \cdot \det(\mathrm{U}) \neq 0$$

在 K 内有两个单根 (即不同的根), 如果条件满足, 则 U 相似于 (§15, 第 5 小节) 一个对角矩阵.

如果 $\mathrm{Tr}(\mathrm{U}) - 4 \cdot \det(\mathrm{U}) = 0$, U 未必相似于一个对角矩阵. 例如

$$\mathrm{U} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

为了 U 相似于 (§15, 第 5 小节) 一个对角矩阵, U 必须在 K^2 具有两个线性无关的即不成比例的特征向量; 而所考虑的矩阵 U 仅有特征值 1, 对应的特征向量是 $x=(\xi_1,\xi_2)$, 使得 $Ux=x$, 即

$$\begin{aligned}\xi_1+\xi_2&=\xi_1,\\ \xi_2&=\xi_2,\end{aligned}$$

即它们是典范基的第一个向量 $e_1=(1,0)$ 的标量倍, 故不可能在 K^2 找到不成比例的两个特征向量, 因此 U 不可能相似于对角矩阵.

例 4 取 $K=\mathbf{R}$ 和 $n=2$, 那么当且仅当

$$\mathrm{Tr}(U)^2-4\cdot\det(U)>0$$

推论可以应用. 在这种情形, 存在一个实可逆矩阵 P, 使得 PUP^{-1} 是对角的. 反之如果有

$$\mathrm{Tr}(U)^2-4\cdot\det(U)<0,$$

存在一个复可逆矩阵 P, 使得 PUP^{-1} 是对角的 (应用例 3), 但是不能够取 P 为实矩阵. 最后假定

$$\mathrm{Tr}(U)^2-4\cdot\det(U)=0,$$

设 $\lambda\in\mathbf{C}$ 是 U 的唯一特征值, 其实它是实数, 即根据判别式为零的二次方程的求根公式有

$$\lambda=\frac{\mathrm{Tr}(U)}{2}.$$

如果存在可逆矩阵 P (实的或复的), 使得

$$PUP^{-1}=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix},$$

那么显然 p_U 也是此式右端的对角矩阵的特征多项式, 即 $(X-\lambda_1)(X-\lambda_2)$. 由于根据假设 p_U 具有重根, 故有

$$\lambda_1=\lambda_2=\lambda,$$

因此

$$PUP^{-1}=\lambda\cdot 1_2.$$

由此得到

$$U=\lambda\cdot 1_2.$$

于是当 $\mathrm{Tr}(U)-4\cdot\det(U)=0$ 时不可能对角化 U, 除非 U 本来就是对角的.

7. 可对角化的自同态的特征

设 u 是交换域 K 上的 n 维向量空间 E 的一个自同态. 我们称 u 是**可对角化的**, 如果存在由 u 的特征向量组成的一个基, 即关于这个基 u 的矩阵是对角矩阵. 定理 5 给了一个矩阵是对角矩阵的充分条件: 即 u 的特征多项式的所有根都在 K 内并且都是单根 (或宁肯说这个多项式在 K 内具有 n 个两两不同的根). 但是这个条件显然不是必要的 (平凡的例子: 单位同态, 它不管关于哪个基的矩阵都是对角的, 但是它的特征多项式即

$$(1-\mathrm{X})^n$$

不具有单根! 假定 $n>1$).

我们要对于一个自同态可对角化陈述一个*充分并且必要的*条件. 为此我们需要下列概念. 给定 u 的一个特征值 λ, 我们称 λ 作为多项式 p_u 的根的重数为 λ 的**重数**. 此外, 我们称相伴于 λ 的特征向量和零向量组成的集合 (这显然是 E 的一个向量子空间) 为**相伴于λ 的 E 的特征子空间**, 用 $\mathrm{E}(\lambda)$ 表示这个子空间. 于是有

$$\mathrm{E}(\lambda)=\mathrm{Ker}(u-\lambda\cdot 1).$$

定理 6 *设 u 是域 K 上的有限维向量空间 E 的一个自同态. u 是可对角化的, 必须并且只需以下两个条件满足:*

a) 特征多项式 p_u 的所有的根都在 K 内;

b) 对于所有特征值 λ, 特征子空间 $\mathrm{E}(\lambda)$ 的维数等于 λ 作为 p_u 的根的重数.

假定这些条件满足, 设 $\lambda_1,\cdots,\lambda_q$ 是 p_u 的两两不同的根, 而 $r_1,\cdots,r_q$ 分别是它们的重数. 根据定理 4, 向量子空间 $\mathrm{E}(\lambda_1),\cdots,\mathrm{E}(\lambda_q)$ 在 §17 第 3 小节的意义下是线性无关的, 由于关系(*)

$$\dim\mathrm{E}(\lambda_i)=r_i,\quad r_1+\cdots+r_q=n,$$

我们还有

$$\dim\mathrm{E}(\lambda_1)+\cdots+\dim\mathrm{E}(\lambda_q)=n=\dim\mathrm{E},$$

由此发现 (§19, 定理 13 的推论 3)

$$\mathrm{E}=\mathrm{E}(\lambda_1)\oplus\cdots\oplus\mathrm{E}(\lambda_q)\ (\text{直和}). \tag{17}$$

因此我们合并 $\mathrm{E}(\lambda_1),\cdots,\mathrm{E}(\lambda_q)$ 的基就得到 E 的一个基, 这个基由 u 的特征向量组成, 从而 u 是可对角化的.

反之, 假定 u 是可对角化的, 仍然用 $\lambda_1,\cdots,\lambda_q$ 表示它的不同的特征值, 并且令

$$s_i=\dim\mathrm{E}(\lambda_i).$$

(*) 第一个表示假设 b), 第二个表示假设 a).

子空间 $E(\lambda_i)$ 是线性无关的 (定理 4). 由于 E 具有由 u 的特征向量即这些子空间 $E(\lambda_i)$ 的元素组成的基, 我们发现这些子空间生成 E. 像上面一样我们有关系 (17), 并且合并 $E(\lambda_1),\cdots,E(\lambda_q)$ 的基就得到 E 的一个基. 关于这个基, u 的矩阵显然是

$$U=\begin{pmatrix}\lambda_1 & & & & & & \\ & \ddots & & & & 0 & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_q & & \\ & 0 & & & & \ddots & \\ & & & & & & \lambda_q\end{pmatrix},$$

其中有 s_i 个对角元素等于 λ_i, U 的也就是 u 的特征多项式是

$$p_u(X)=(\lambda_1-X)^{s_1}\cdots(\lambda_q-X)^{s_q}.$$

这就证明了: 一方面它的特征值全部在 K 内, 另一方面 λ_i 的重数等于 $s_i=\dim E(\lambda_i)$, 这就完成了证明.



注 3 如果 K 是代数闭的, 条件 a) 总是满足的, 例如 $K=\mathbf{C}$.

注 4 条件 b) 意即 $E(\lambda)$ 具有最大可能的维数 (这符合找到充分多的特征向量以便从中抽取 E 的一个基的美好愿望); 换句话说, 对于任意自同态 u 和 u 在 K 内的所有特征值有不等式

$$\dim E(\lambda)\leqslant\lambda\text{ 的重数}.$$

为了证明这一点, 令 $E(\lambda)=F$. 显然 $u(F)\subset F$, 那么由 u 限制到 F 的自同态 v 是比例为 λ 的位似, 故 v 的特征多项式是

$$p_v(X)=(\lambda-X)^{\dim E(\lambda)},$$

于是, 为了证明所要的结果, 需要验证 $p_v(X)$ 整除 $p_u(X)$. 换句话说, 这就归结为建立一般的结果, 它就是: *设 F 是 E 的在 u 下是稳定的一个线性子空间, 即 $u(F)\subset F$, 设 v 是由 u 限制到 F 的自同态, 即对于 $x\in F$ 有 $v(x)=u(x)$. 那么多项式 $p_v(X)$ 整除多项式 $p_u(X)$.*

为此, 选择 F 的任意一个基 $x_1,\cdots,x_r$, 并且补充成 E 的一个基 (这总是可以的: §19, 定理 2 的推论 2). u 关于这个基的矩阵显然有形式

$$U=\begin{pmatrix}V & T\\ 0 & W\end{pmatrix},$$

其中 V 是 v 关于基 $x_1, \cdots, x_r$ 的矩阵, T 是 r 行 $n-r$ 列矩阵, 而 W 是 $n-r$ 阶方阵, 因此有

$$\mathrm{U} - \mathrm{X}\cdot 1_n = \begin{pmatrix} \mathrm{V} - \mathrm{X}\cdot 1_r & \mathrm{T} \\ 0 & \mathrm{W} - \mathrm{X}\cdot 1_{n-r} \end{pmatrix},$$

而 §24 例 2 指出

$$\det(\mathrm{U} - \mathrm{X}\cdot 1_n) = \det(\mathrm{V} - \mathrm{X}\cdot 1_r)\det(\mathrm{W} - \mathrm{X}\cdot 1_{n-r}),$$

此式可以改写为

$$p_u(\mathrm{X}) = p_v(\mathrm{X})p_{\mathrm{W}}(\mathrm{X}),$$

这就证明了所要的结果.

可对角化矩阵的概念对应于可对角化自同态概念. 说一个矩阵 $\mathrm{U} \in \mathrm{M}_n(\mathrm{K})$ 是**可对角化的**, 如果存在一个矩阵 $\mathrm{P} \in \mathrm{GL}(n, \mathrm{K})$, 使得 PUP^{-1} 是对角的. 基于 §15 第 5 小节的理由, 其意义就是说由 U 定义的 K^n 的自同态是可对角化的. 如果特征多项式 p_{U} 的所有根都在 K 内并且它们都是单根, 则 U 总是可对角化的.

一个类似但稍微精妙的概念是**半单矩阵**的概念. 称具有下列性质的所有矩阵 $\mathrm{U} \in \mathrm{M}_n(\mathrm{K})$ 是半单的: 存在 K 的一个扩张 L, 使得 U 作为元素在 L 内的矩阵是可对角化的 (人们证明了这时可以取 K 的不论哪一个代数闭扩张作为 L). 如果把 U 看作 K^n 的一个自同态 u 的矩阵, 那么直观地这就是说 u 具有 "足够多的" 特征向量, 只要允许其坐标在 K 的一个扩张域内.

例 5 取 $\mathrm{K} = \mathbf{R}$ 和矩阵

$$\mathrm{U} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

它在 $\mathbf{C}$ 内的特征值 (注 1) 是

$$\cos\theta \pm i\cdot\sin\theta.$$

如果 θ 不是 π 的整倍数, 那么这些特征值是不同的, 因此看作元素在 $\mathbf{C}$ 内的矩阵 U 是可对角化的; 但是 U 看作元素在 $\mathbf{R}$ 内的矩阵它不是可对角化的, 因为它的特征值不是实数. 即在域 $\mathbf{R}$ 上, 矩阵 U 仅是半单的 (其中显然包括 θ 是 π 的整倍数的情形).

§34 习题

对于出现在习题 1 直至 14 中的每一个矩阵, 解答下列问题:

a) 计算特征值 (取 $\mathbf{C}$ 作为基础域).

b) 对于每一个特征值计算对应的在 $\mathbf{C}^n$ 内的特征向量 (把每一个复元素的 n 阶方阵等同于 $\mathbf{C}^n$ 的一个自同态) .

c) 求由特征向量组成的 $\mathbf{C}^n$ 的一个基, 如果这种情形出现.

d) 如果所考虑的矩阵在 $\mathbf{C}$ 上是可对角化的, 确定 $\mathbf{C}$ 的最小子域, 在该子域上矩阵是可对角化的.

e) 如果所考虑的矩阵在 $\mathbf{C}$ 上不是可对角化的, 求 $\mathbf{C}^n$ 的一个基, 使得对应的自同态关于这组基具有三角矩阵.

1. $\begin{pmatrix} 5 & -3 & 2 \\ 6 & -4 & 4 \\ 4 & -4 & 5 \end{pmatrix}$.

2. $\begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}$.

3. $\begin{pmatrix} 4 & -5 & 7 \\ 1 & -4 & 9 \\ -4 & 0 & 5 \end{pmatrix}$.

4. $\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$.

5. $\begin{pmatrix} 9 & -6 & -2 \\ 18 & -12 & -3 \\ 18 & -9 & -6 \end{pmatrix}$.

6. $\begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$.

7. $\begin{pmatrix} 4 & 6 & -15 \\ 3 & 4 & -12 \\ 2 & 3 & -8 \end{pmatrix}$.

8. $\begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}$.

9. $\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}$.

10. $\begin{pmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix}$.

11. $\begin{pmatrix} 0 & 0 & 2 & 3 \\ 0 & 0 & -2 & -3 \\ 2 & -2 & 0 & -1 \\ 3 & -3 & -1 & -3 \end{pmatrix}$.

12. $\begin{pmatrix} 3 & 2 & 1 & -1 \\ 2 & 2 & 1 & -1 \\ 1 & 1 & 1 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix}$.

13. $\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 3 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & 2 & 3 & 4 & \cdots & n \end{pmatrix}$.

14. $\begin{pmatrix} 0 & c & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & c & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & c \\ c & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$ (n 行 n 列).

15. 设 L 是交换域 K 上的有限维向量空间, 而 u 是 L 第一个自同态. 设 L^* 是 L 的对偶空间, 而 u 的转置 ${}^t u$ 是 L^* 的自同态.

a) 证明 u 的特征值和 ${}^t u$ 的特征值是相同的.

b) 设 $\lambda \in \mathrm{K}$ 是 u 的特征值. 用 $\mathrm{E}(\lambda)$ 表示使得 $u(x) = \lambda x$ 的 $x \in \mathrm{L}$ 的集合, 而 $\mathrm{F}(\lambda)$ 表示使得 ${}^t u(f) = \lambda f$ 的 $f \in \mathrm{L}^*$ 的集合. 证明

$$\dim \mathrm{E}(\lambda) = \dim \mathrm{F}(\lambda)$$

(利用 §19, 习题 12).

c) 证明 u 的特征多项式和 ${}^t u$ 的特征多项式是同样的. (事实上, 可以证明其元素在一个交换域内 K 的所有方阵相似于它的转置矩阵, 参见 §35, 习题 10).

16. 设 $\mathrm{G} = \mathrm{SL}(2, \mathbf{R})$ 是实元素的并且其行列式

$$ad - bc = 1$$

的矩阵

$$\mathrm{X} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

的群.

a) 如果

$$|\mathrm{Tr}(\mathrm{X})| > 2,$$

则存在矩阵 $\mathrm{U} \in \mathrm{G}$, 使得

$$\mathrm{UXU}^{-1} = \begin{pmatrix} t & 0 \\ 0 & 1/t \end{pmatrix},$$

其中的 $t \in \mathbf{R}, t \neq 0, 1, -1$ (这时称 X 是**双曲的**).

b) 如果

$$|\mathrm{Tr}(\mathrm{X})| < 2,$$

则存在矩阵 $\mathrm{U} \in \mathrm{G}$, 使得

$$\mathrm{UXU}^{-1} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix},$$

其中的 $t \in \mathbf{R}, t \neq 0, 1, -1$ (这时称 X 是**椭圆的**).

c) 设 $\mathrm{X} \neq 1_2, -1_2$, 并且

$$|\mathrm{Tr}(\mathrm{X})| = 2,$$

则存在矩阵 $\mathrm{U} \in \mathrm{G}$, 使得 UXU^{-1} 等于下列矩阵之一:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix}.$$

(这时称 X 是**抛物的**.)

d) 用 K 表示由形如

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

(t 为实数) 的矩阵组成的 G 的子群, 用 T 表示形如

$$\begin{pmatrix} u & v \\ 0 & 1/u \end{pmatrix}$$

的矩阵组成的 G 的子群, 其中的 u, v 是实数, 并且 $u > 0$. 证明 G 的所有元素以唯一的一种方式写成形式 XY, $\mathrm{X} \in \mathrm{K}, \mathrm{Y} \in \mathrm{T}$. 由此推出在问题 a) 里, 可以假定 $\mathrm{U} \in \mathrm{K}$, 而在问题 b) 里, 可以假定 $\mathrm{U} \in \mathrm{T}$.

17. 设 $G = SL(2, \mathbf{C})$ 是复元素的并且其行列式

$$ad - bc = 1$$

的矩阵

$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

的群.

a) 如果

$$\mathrm{Tr}(X) \neq 2, -2,$$

则存在矩阵 $U \in G$, 使得

$$UXU^{-1} = \begin{pmatrix} t & 0 \\ 0 & 1/t \end{pmatrix},$$

其中的 $t \in \mathbf{C}, t \neq 0, 1, -1$.

b) 证明, 如果 $X \neq 1_2, -1_2$, 并且

$$\mathrm{Tr}(X) = 2 \text{ (对应的, } -2),$$

则存在矩阵 $U \in G$, 使得

$$UXU^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ (对应的, } \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}).$$

c) 用 K 表示形如

$$\begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix}$$

的矩阵的集合, 其中 u 和 v 是复数, 并且满足关系

$$|u|^2 + |v|^2 = 1,$$

用 T 表示 G 中的矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 的集合, 其中 $c = 0$, 并且 a 和 d 是正实数. 证明 K 和 T 是 G 的子群, 并且 G 的所有元素以唯一的一种方式表示成 XY, $X \in K$ 并且 $Y \in T$.

¶18. 设 K 是一个交换环. 给定一个系数在 K 内的多项式

$$f(X) = a_0 + a_1X + \cdots + a_rX^r,$$

对于所有矩阵 $A \in M_n(K)$ 定义

$$f(A) = a_0 \cdot 1_n + a_1A + \cdots + a_rA^r.$$

a) 设 f 和 g 是系数在 K 内的两个多项式, 令

$$f + g = p, \quad fg = q, \quad f(g(X)) = r(X).$$

证明

$$p(\mathrm{A}) = f(\mathrm{A}) + g(\mathrm{A}), \quad q(\mathrm{A}) = f(\mathrm{A})g(\mathrm{A}) = g(\mathrm{A})f(\mathrm{A}), \quad r(\mathrm{A}) = f(g(\mathrm{A})).$$

(用恰当的方式利用 §28 的结果, 尽量避免任何计算.)

b) 证明如果 $\mathrm{A} \in \mathrm{M}_n(\mathrm{K})$ 并且 $\mathrm{U} \in \mathrm{GL}(n, \mathrm{K})$, 则有

$$f(\mathrm{UAU}^{-1}) = \mathrm{U}f(\mathrm{A})\mathrm{U}^{-1}.$$

c) 假定 A 是三角矩阵, 并且用 $t_1, \cdots, t_n$ 表示它的对角线上的元素. 证明 $f(\mathrm{A})$ 是三角矩阵, 并且它的对角线上的元素是 $f(t_1), \cdots, f(t_n)$.

d) 假定 K 是一个域. 设 $t_1, \cdots, t_n$ 是 $\mathrm{A} \in \mathrm{M}_n(\mathrm{K})$ 的特征值 (在 K 的一个代数闭扩张内取的, 并且每一个特征值重复的次数等于作为 A 的特征方程的根的重数, 读者如果对于一般情形不感兴趣, 可以假定 $\mathrm{K} = \mathbf{C}$). 证明 $f(\mathrm{A})$ 的特征值是 $f(t_1), \cdots, f(t_n)$, 并且

$$\det(f(\mathrm{A})) = f(t_1) \cdots f(t_n), \quad \mathrm{Tr}(f(\mathrm{A})) = f(t_1) + \cdots + f(t_n).$$

¶19. 设

$$f(\mathrm{X}) = a_1 + a_2\mathrm{X} + \cdots + a_n\mathrm{X}^{n-1}$$

是一个复系数多项式. 证明

$$\begin{vmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{vmatrix} = f(z_1) \cdots f(z_n),$$

其中 $z_1, \cdots, z_n$ 是 $\mathbf{C}$ 中的单位的 n 次根 (这样的行列式称为**轮换行列式**; 利用前一个习题).

利用这个结果计算行列式

$$\begin{vmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{vmatrix}, \quad \begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \\ 4 & 5 & 1 & 2 & 3 \\ 5 & 1 & 2 & 3 & 4 \end{vmatrix}.$$

20. 设 s 是整数 $1, 2, \cdots, n$ 的一个置换. 考虑 $\mathbf{C}^n$ 的由

$$u_s(e_i) = e_{s(i)} \quad (1 \leqslant i \leqslant n)$$

给定的一个自同态 u_s, 这里 $(e_i)_{1 \leqslant i \leqslant n}$ 是 $\mathbf{C}^n$ 的典范基. 利用 s 的轮换分解 (§7, 习题 24) 计算 u_s 的特征值, 并且证明 u_s 是可对角化的.

¶将前面的 $\mathbf{C}$ 换成一个特征 $p \neq 0$ 的代数闭域 K. 取 $n = p$, 取一个轮换为 s. 证明 u_s 不是可对角化的.

21. 设 V 是交换域 K 上的一个有限维向量空间, 而 F 是 V 的自同态的一个集合. 称 F 是**可三角化的** (或对于 F **同时化成三角形矩阵**是可能的), 如果存在 V 的一个基, 使得所有 $u \in \mathrm{F}$ 关于这个基的矩阵都是三角矩阵.

设 V′ 关于 F 是 V 的**稳定子空间**, 即有

$$u(\mathrm{V}') \subset \mathrm{V}' \quad \text{对于所有 } u \in \mathrm{F}$$

(人们还用*不变的*来代替稳定的). 所有的 $u \in \mathrm{F}$ 诱导 V′ 的一个自同态 u' 和商向量空间 V/V' 的一个自同态 $\bar{u}$ (§10, 习题 10). 设 F′ 是 u' 的集合, 而 $\bar{\mathrm{F}}$ 是 $\bar{u}$ 的集合, u 遍历 F.

假定在 V′ 内 F′ 是可三角化的, 并且在 V/V' 内 $\overline{\mathrm{F}}$ 是可三角化的. 证明 F 在 V 内是可三角化的 (仿照 §18 的定理 1 的证明构造在 V 内的一个基).

由此得到 §34 的定理 3 的证明.

22. 设 V 是在代数闭域 K 上的一个有限维向量空间 (例如, 可以假定 $\mathrm{K} = \mathbf{C}$, 证明是同样的), 而 F 是 V 的两两交换的自同态的集合. 我们打算证明 F 是可三角化的 (见上题).

a) 对于所有的 $u \in \mathrm{F}$ 和 u 的所有特征值 λ, 设 $\mathrm{V}_u(\lambda)$ 是使得 $u(x) = \lambda x$ 的 $x \in \mathrm{V}$ 的子空间. 证明 $\mathrm{V}_u(\lambda)$ 在 F 下是稳定的.

b) 由此推出各个 $u \in \mathrm{F}$ 在 V 内有公共的特征值 [利用问题 a) 并关于 $\dim(\mathrm{V})$ 进行归纳推理].

c) 利用前一个习题结束证明.

d) 此外假定每一个 $u \in \mathrm{F}$ 是可对角化的. 证明存在 V 的一个基, 使得所有 $u \in \mathrm{F}$ 关于这个基的矩阵是对角的 (两两交换的可对角化的自同态 "**同时化成对角形**").

e) 证明代替假定 K 是代数闭的, 只需假定所有的 $u \in \mathrm{F}$ 是可三角化的 [或对于问题 d), 可对角化的].

¶23. 设 V 是交换域 K 上的有限维向量空间, 而 F 是 V 的自同态的一个集合. 称 F 是**不可约的**, 如果 V 在 F 下稳定的子空间仅有 $\{0\}$ 和 V.

a) 证明, 如果 V 的一个自同态与所有 $u \in \mathrm{F}$ 交换, 则向量子空间 $\mathrm{Ker}(f)$ 和 $\mathrm{Im}(f)$ 以及 f 的特征子空间在 F 下是稳定的.

b) (Schur 引理) 假定 F 是不可约的, 并且 K 是代数闭的. 证明与所有 $u \in \mathrm{F}$ 交换的 V 的自同态仅是位似自同态.

c) 仍然假定 F 是不可约的, 但是关于 K 不再做假设. 证明与所有 $u \in \mathrm{F}$ 交换的 V 的自同态组成 V 的自同态环的一个子域 (可能是不交换的).

d) 取 $\mathrm{K} = \mathbf{R}$ 和 $\mathrm{V} = \mathbf{R}^4$. 选择 F, 使得上一个问题中的子域是 §15 习题 11 的四元数域.

¶24. 设 V 是特征为 0 的交换域 K 上的 n 维向量空间, 而 G 是 V 的自同构的一个*有限群*, 用 r 记 G 的阶.

a) 设 f 是 V 的一个同态, 证明自同态

$$f^{\#} = \frac{1}{r} \sum_{s \in \mathrm{G}} S \circ f \circ S^{-1}$$

同所有 $s \in \mathrm{G}$ 交换, 并且当且仅当 f 同 G 的元素 s 交换有 $f^{\#} = f$. 证明如果 g 同 G 的所有元素 s 交换, 则有

$$f \circ g^{\#} = f^{\#} \circ g.$$

b) 设 W 是在 G 下不变的 V 的向量子空间, 即 (习题 21) 对于所有 $s \in G$ 有 $s(W) \subset W$ (我们将顺便证明其实有对于所有 $s \in G$ 有 $s(W) = W$). 选择 (§17 定理 2 的推论, 结合 W 具有一个在 V 内的补空间这个事实) V 的一个自同态 p, 使得

$$p^2 = p, \quad p(V) = W.$$

证明

$$\mathrm{Im}(p^{\#}) = W.$$

c) 通过考虑在问题 b) 里 $p^{\#}$ 的核, 证明下列定理: 所有在 G 下不变的 V 的子空间在 V 内具有一个在 G 下不变的补空间.

d) 设 V 是特征为 0 的代数闭域上的有限维向量空间 (例如 **C**) , 而 G 是 V 的自同构的一个交换群. 证明存在 V 的一个基, 使得所有 $s \in G$ 关于这个基的矩阵是对角的; 此外如果 G 是 n 阶的, 则这些矩阵的对角线上的元素是单位的 n 次根 [利用习题 22 的 b)].

e) 设 X 是元素在特征为 0 的代数闭域内的一个矩阵, 如果对于一个整数 $n \geqslant 1$ 有

$$X^n = 1,$$

则 X 是可对角化的. 通过一个例子证明这个结果不能推广到特征 $p \neq 0$ 的域.

f) 证明问题 c) 的结果对于特征 $p \neq 0$ 的域仍然有效, 只要群 G 的阶不是 p 的倍数. 对于问题 d) 有同样的结果.

¶25. 设 V 是特征为 0 的代数闭域上的 $n+1$ 维向量空间. 考虑 V 的满足下列交换公式

$$[h, u] = 2u, \quad [h, v] = -2v, \quad [u, v] = h$$

的三个自同态 u, v 和 h, 其中按照一般定义 $[f, g] = f \circ g - g \circ f$. 最后假定集合 $\{u, v, h\}$ 是不可约的, 即 (习题 23) 同时在 u, v 和 h 下稳定的向量子空间仅是 $\{0\}$ 和 V.

a) 设 $x \in V$ 和 $\lambda \in K$ 使得 $h(x) = \lambda x$. 证明向量 $y = u(x)$ 满足 $h(y) = (\lambda + 2)y$, 并且 $z = v(x)$ 满足 $h(z) = (\lambda - 2)z$.

b) 证明存在一个向量 $x \neq 0$ 和一个标量 $\lambda \in K$, 使得

$$h(x) = \lambda x, \quad u(x) = 0.$$

c) x 是满足 b) 的条件的向量, 令

$$x_k = v^k(x)/k! \quad (k \geqslant 0).$$

证明关系

$$\begin{aligned} h(x_k) &= (\lambda - 2k)x_k, \\ u(x_k) &= (\lambda - k + 1)x_{k-1}, \\ v(x_k) &= (k+1)x_{k+1}. \end{aligned}$$

d) 考虑到上面所做的不可约假设, 由这些关系推出 $\lambda = n$, 并且 $n+1$ 个向量 $x_0, x_1, \cdots, x_n$ 组成 V 的一个基. u, v 和 h 关于这个基的矩阵是什么? 情形 $n = 2$ 或 3 怎样?

e) 取系数在 K 内的一个未定元的至多 n 次的多项式组成的向量空间为 V. 证明前一个问题所描述的情况事实上会实现, 如果如下定义 u, v 和 h: u 把每一个多项式 $f(\mathrm{X})$ 变换成多项式 $n\mathrm{X}f(\mathrm{X}) - \mathrm{X}^2f'(\mathrm{X})$, v 把每一个多项式 $f(\mathrm{X})$ 变换成多项式 $f'(\mathrm{X})$, 而 h 变换 $f(\mathrm{X})$ 成 $-nf(\mathrm{X}) + 2\mathrm{X}f'(\mathrm{X})$. 自然 f' 表示的是 f 的导多项式.

¶26. 设 V 是特征为 0 的代数闭域 K 上的有限维向量空间 (例如 $\mathrm{K} = \mathbf{C}$), 而 u, v, w 是 V 的三个自同态. 假定

$$[u, w] = [v, w] = 0, \quad [u, v] = w.$$

证明存在 V 的一个基, 使得 u, v, w 关于这个基的矩阵是三角矩阵. 当 V 的维数是 3 时确定问题的解 u, v, w.

¶¶27. 设 V 是交换域 K 上的有限维向量空间. V 的自同态的一个集合 F 称为 (V 的自同态的) 一个 **Lie 代数**, 如果 F 是一个向量空间 (即对于任意 $u, v \in \mathrm{F}$ 和 $\alpha, \beta \in \mathrm{K}$, $\alpha u + \beta v \in \mathrm{F}$), 并且还有

$$u \circ v - v \circ u \in \mathrm{F} \quad \text{对于任意 } u, v \in \mathrm{F}.$$

(例子: 取习题 25 的 u, v, h 或习题 26 的 u, v 和 w 的所有线性组合的集合.)

称 V 的自同态的一个 Lie 代数**是可解的**, 如果存在 F 的向量子空间的一个递增序列

$$\{0\} = \mathrm{F}_0 \subset \mathrm{F}_1 \subset \cdots \subset \mathrm{F}_n = \mathrm{F}, \tag{$*$}$$

使得对于满足 $1 \leqslant i \leqslant n$ 的所有 i 有

$$u \circ v - v \circ u \in \mathrm{F}_{i-1}, \quad \text{对于任意 } u, v \in \mathrm{F}_i. \tag{$**$}$$

我们打算证明: *如果 K 是特征为 0 的和代数闭的 (例如 $\mathrm{K} = \mathbf{C}$), 并且 F 是可解的, 则存在 V 的一个基, 使得所有 $u \in \mathrm{F}$ 关于这个基的矩阵是三角矩阵* [Lie 定理, 容易看到它推广了习题 22 的 c) 以及习题 26 的结果].

a) 在序列 $(*)$ 中 $n = 1$ 的情形证明定理.

b) 证明 $(*)$ 的项 F_{i-1} 是可解的一个 Lie 代数.

c) 假定找到 V 内的一个向量 $x \neq 0$, 使得有形式如下的关系:

$$u(x) = \lambda(u)x \quad \text{对于所有 } u \in \mathrm{F}_{i-1} \tag{$***$}$$

(即 x 是各个 $u \in \mathrm{F}_{i-1}$ 的公共特征向量). 取 $v \in \mathrm{F}_n$, 并且令 $y = v(x)$. 证明有

$$u(y) = \mu(u) \cdot y \quad \text{对于所有 } u \in \mathrm{F}_{i-1},$$

其中 $\mu(u)$ 是要计算的标量. 在所得到的结果中用 ξv 代替 v (ξ 是 K 的任意元素), 导出

$$\mu(u) = \lambda(u) \quad \text{对所有的 } u \in \mathrm{F}_{n-1}.$$

d) 用 $\mathrm{V}(\lambda)$ 表示满足 $(***)$ 的 $x \in \mathrm{V}$ 组成的 V 的子空间. 证明在所有 $v \in \mathrm{F}_n$ 下它是稳定的, 并且 F_n 的任意两个元素在 $\mathrm{V}(\lambda)$ 上的限制是交换的 (利用 §§12, 13, 14 的习题 8). 由此推出各个 $u \in \mathrm{F}$ 在 $\mathrm{V}(\lambda)$ 内至少有一个公共特征向量.

e) 关于 V 的维数进行归纳推理以完成证明 (利用习题 21).

f) 域 K 的特征为 0 这个假设用在哪里?

g) 证明 (在对于 K 所做的假设下) Lie 的定理刻画了可解的 Lie 代数的特征.

¶28. 元素在代数闭域 K 内的方阵是幂零的, 必须并且只需它在 K 内的所有的特征值是零.

证明如果 K 的特征是 0 (例如 $K = \mathbf{C}$), 可以把这个条件换为

$$\mathrm{Tr}(X) = \mathrm{Tr}(X^2) = \cdots = \mathrm{Tr}(X^n) = 0,$$

其中的 n 是 X 的阶.

系数在 K 内的矩阵是幂幺的 (即 $1 - U$ 是幂零的), 必须并且只需 U 的仅有的特征值是 1. U 的特征多项式是什么?

证明如果 K 的特征 $p \neq 0$, 可以把这个条件换成下列条件: 存在一个整数 $n \geqslant 0$, 使得

$$U^{p^n} = 1.$$

29. 设 A 是一个可逆的方阵, 其元素在一个代数闭的域内. 证明 A 的逆矩阵的特征值是 A 的特征值的逆, 并且重数相同.

¶30. 设 A 是一个 n 阶可逆的方阵, 其元素在一个交换域 K 内. 设 f 是系数在 K 内的一个未定元的有理分式. 称 A 是**可代入** f 内的, 如果存在多项式 p 和 q, 使得

$$f = p/q, \quad \text{并且} \ \det(q(A)) \neq 0.$$

证明这时矩阵

$$f(A) = p(A) \cdot q(A)^{-1}$$

不依赖 p 和 q 的选取 (只要 p 和 q 满足所断言的条件).

假定 K 是代数闭的, 并且 $\lambda_1, \cdots, \lambda_n$ 是 A 的特征值 (考虑到它们的重数). 证明 A 是可代入 f 内的, 必须并且只需 f 在每一个 λ_i 有定义, 并且 $f(A)$ 的特征值是 $f(\lambda_1), \cdots, f(\lambda_n)$ (利用习题 18).

31. 设 V 是交换域 K 上的有限维向量空间, u 是 V 的一个自同态, 而 W 是在 u 下稳定的 V 的一个向量子空间. 证明如果 u 是可对角化的 (对应的, 可三角化的), 则 u 在 W 内的限制也是可对角化的 (对应的, 可三角化的).

¶32. 设 u 和 v 是交换域 K 上的有限维向量空间 V 的两个自同态. 假定 u 和 v 是可对角化的并且是交换的. 证明 $v \circ u$ 是可对角化的.

¶¶33. 设 K 是一个交换域. 给定一个矩阵 $U \in M_n(K)$, 考虑映射 f_U: $M_n(K) \to M_n(K)$, 其定义是

$$f_U(X) = UX - XU = [U, X] \quad \text{对于所有} \ X \in M_n(K),$$

并且把 f_U 看作向量空间 $M_n(K)$ 的一个自同态. 证明 f_U 是可对角化的, 必须并且只需 U 是可对角化的.

34. 设 K 是一个交换域, 而 n 是一个整数. 证明, 作为 K 上的一个向量空间, 环 $M_n(K)$ 仅有一个由具有下列性质的矩阵组成的一个基: 对于所有的对角矩阵 $H \in M_n(K)$, $[H, X] = \alpha(H) \cdot X$, 这里 $\alpha(H)$ 是一个依赖于 H 的标量, 计算它.

¶35. 设 V 是交换域 K 上的一个有限维向量空间, 用 $T^p_q(V)$ 表示 V 上的 p 次共变 q 次反变的张量的向量空间 (§21, 例 6). 设 u 是 V 的一个自同构, 考虑 §21 习题 1 中定义的 $T^p_q(V)$ 的自

同构 $T_q^p(u)$. 证明如果 u 是可对角化的则 $T_q^p(u)$ 也是可对角化的. 还证明如果 V 的一个自同态 u 是可对角化的, 则 §21 习题 1 中定义的 $T_q^p(V)$ 的自同态 $D_q^p(u)$ 也是可对角化的. 在这两种情形的每一个中, 通过 u 的特征值计算 $T_q^p(V)$ 的所考虑的自同态的特征值.

¶36. 保留上个习题的记号, 选择 V 的一个基 $(a_i)_{1\leqslant i\leqslant n}$, 并且用 G 表示关于这个基的矩阵为三角矩阵 (对应的, 对角矩阵) 的 V 的自同构的群. 构造 $T_q^p(V)$ 的一个基, 使得对于每一个 $u\in G$, $T_q^p(u)$ 关于这个基的矩阵是三角矩阵 (对应的, 对角矩阵) .

¶37. 设 V 是交换域 K 上的一个有限维向量空间, 而 $S_r(V)$ 是 V 上的 r 次齐次多项式组成的向量空间 (§§27, 28, 习题 17). 令 V 的每一个自同构 u 对应 $S_r(V)$ 的由 $u_r(f)=f\circ u$ 定义的自同构 u_r.

证明如果 u 是可对角化的 (对应的, 可三角化的), 则 u_r 也是. 用 u 的特征值计算 u_r 的特征值.

假定 u 是可对角化的, 用 u 的特征多项式的系数计算 $\mathrm{Tr}(u_3)$. 得到的结果能够推广到 V 的所有自同构吗? 对于任意 r 是否存在计算 $\mathrm{Tr}(u_r)$ 的类似公式?

38. 不利用特征值是一个代数方程的根这个事实证明 §34 的定理 4 (写出 $x_1,\cdots,x_n$ 之间的非平凡线性关系, 把 u 作用于这个关系, 由此得到向量 $x_1,\cdots,x_n$ 中的 $n-1$ 个向量之间的一个非平凡线性关系).

39. 利用 Vandermonde 行列式 (§24, 习题 14) 证明 §34 的定理 4 (写出 $x_1,\cdots,x_n$ 之间的非平凡线性关系, 把 $u,u^2,\cdots,u^{n-1}$ 依次作用于这个关系).

¶40. 设 $A=(a_{ij})_{1\leqslant i,j\leqslant n}$ 是元素在一个环 K′ 内的 n 阶方阵. 给定集合 $\{1,2,\cdots,n\}$ 的两个子集 H 和 K, 用 $A_{H,K}$ 表示由 A 的元素 a_{ij} 组成的矩阵, 其中的 $i\in H, j\in K$. 设

$$(-1)^n p_A(X)=X^n-\tau_1(A)X^{n-1}+\tau_2(A)X^{n-2}-\cdots$$

(§34, 第 3 小节, 公式 (9)). 证明这个多项式的系数 $\tau_p(A)$ 由公式

$$\tau_p(A)=\sum\det(A_{H,H})$$

给定, 其中的和取遍 $\{1,2,\cdots,n\}$ 的所有满足条件 $\mathrm{Card}(H)=p$ 的子集 H 上.

¶¶41. 设 L 是一个交换环, 而 A 是 L 的一个子环. 称 $x\in L$ 在 A 上是**整的**, 如果它满足形如

$$x^n+a_{n-1}x^{n-1}+\cdots+a_0=0$$

的方程, 其中的 $a_{n-1},\cdots,a_0\in A$ (如果 L 不是一个域, 那么 x^n 的系数是 1 是本质的). 一个复数称为**代数整数**, 如果它在 $\mathbf{C}$ 的子环 $\mathbf{Z}$ 上是整的.

a) 假定前面的 L 作为 A-模是有限生成的. 设 $(m_i)_{1\leqslant i\leqslant r}$ 是 A-模 L 的一个生成元组. 证明对于所有 $x\in L$, 存在 $a_{ij}\in A$, 使得

$$xm_i=\sum_{j=1}^r a_{ij}m_j\quad(1\leqslant i\leqslant r).$$

令

$$\begin{vmatrix} a_{11}-x & a_{12} & \cdots & a_{1r}\\ a_{21} & a_{22}-x & \cdots & a_{2r}\\ \vdots & \vdots & & \vdots\\ a_{r1} & a_{r2} & \cdots & a_{rr}-x \end{vmatrix}=d,$$

证明对于 $1 \leqslant i \leqslant r$ 有 $dm_i = 0$. 由此推出 $d = 0$, 于是所有 $x \in \mathrm{L}$ 在 A 上是整的.

b) 不再假定 L 作为 A-模是有限生成的, 设 $x_1, \cdots, x_q$ 是 L 的在 A 上是整的元素, 证明 $\mathrm{A}[x_1, \cdots, x_q]$ 是一个有限生成的 A-模.

c) 证明在 A 上是整的 $x \in \mathrm{L}$ 的集合 B 是 L 的一个子环 (称为 A 在 L 内的**整闭包**). 例子: 代数整数组成 $\mathbf{C}$ 的一个子环. [这个结果 (以及上面的推理) 属于 Dedekind.]

d) 设 C 是一个交换环, B 是 C 的一个子环, 而 A 是 B 的一个子环. 假定所有 $x \in \mathrm{C}$ 在 B 上是整的, 并且所有 $x \in \mathrm{B}$ 在 A 上是整的. 证明所有 $x \in \mathrm{C}$ 在 A 上是整的.

¶42. 元素为有理整数的一个方阵的特征值是代数整数. 反之, 所有代数整数是一个元素在 $\mathbf{Z}$ 内的一个方阵的特征值.

43. 所有的是代数整数的有理数是有理整数.

¶44. 考虑有理数域的二次扩张

$$\mathrm{L} = \mathbf{Q}[\sqrt{d}].$$

假定 d 是一个有理整数, 并且不被任何素数的平方整除 (顺便将证明总可以把 $\mathbf{Q}$ 的二次扩张归结为这种情形).

a) L 的一个元素 $z = x + y\sqrt{d}(x, y \in \mathbf{Q})$ 在 $\mathbf{Z}$ 上是整的, 必须并且只需

$$2x \in \mathbf{Z} \quad \text{并且} \quad x^2 - y^2 d \in \mathbf{Z}$$

(注意到如果 z 是一个代数整数, 则 $\bar{z} = x - y\sqrt{d}$ 亦然).

b) 设 B 是所有在 $\mathbf{Z}$ 上是整数的 $z \in \mathrm{L}$ 的环. 证明, 作为加法群的 L 具有由两个元素

$$1, \sqrt{d}, \quad \text{如果 } d \equiv 2 \text{ 或 } 3 (\mathrm{mod}\ 4),$$
$$1, \frac{1+\sqrt{d}}{2}, \quad \text{如果 } d \equiv 1\ (\mathrm{mod}\ 4)$$

组成的基.

45. 证明所有代数数是一个代数整数除以一个非零有理整数的商.

¶¶46. 称一个交换整环 A 是**整闭的**, 如果 A 的分式域的在 A 上是整的所有元素属于 A. 例子: 环 $\mathbf{Z}$ (习题 43).

a) 假定 A 是一个赋值环 (即对于所有 $x \in \mathrm{K}$ 有 $x \in \mathrm{A}$ 或 $x^{-1} \in \mathrm{A}$, 参见 §8, 习题 6). 证明 A 是整闭的.

b) 如果 A 是它的分式域 K 的赋值环的交集, 则 A 是整闭的 [注意: 可以证明其逆].

c) 所有唯一因式分解整环 (§31, 习题 21) 是整闭的, 同样所有 Dedekind 整环 (§§10, 11, 习题 14 和 §18, 习题 7) 也如此.

d) 如果 A 是整闭的, 并且如果 $\mathfrak{p}$ 是 A 的一个素理想, 则局部环 $\mathrm{A}_{\mathfrak{p}}$ [§29, 习题 9, e): $\mathrm{A}_{\mathfrak{p}}$ 是 $x \in \mathrm{K}$ 的集合, 它可以写成形式 u/v, 其中的 $u, v \in \mathrm{A}$ 并且 $v \notin \mathfrak{p}$] 是整闭的.

e) 设 A 是一个交换整环, K 是它的分式域, 而 L 是 K 的一个扩张, 那么 A 在 L 内的整闭包是一个整闭环.

¶¶47. 设 A 是一个整闭环, 而 K 是它的分式域.

a) 设 E 是 K 的一个代数闭扩张, 而 x 是 E 的一个在 A 上是整的元素 (从而在 K 上是代数的). 设 f 是 x 在 K 上的极小多项式 (§32, 习题 9 和 10). 证明 f 在 E 内的所有的根在 A 上是

整的. 由此得到结论: f 的系数属于 A. (因此, 为了验证在 K 上是代数的一个元素 x 在 A 上是整的, 只需考察它在 K 上的极小方程.)

b) 设 L 是 K 的有限扩张. 证明对于在 A 上是整的所有 $x \in \mathrm{L}$ 有

$$\mathrm{Tr}_{\mathrm{L/K}}(x) \in \mathrm{A}, \quad \mathrm{N}_{\mathrm{L/K}}(x) \in \mathrm{A}$$

(§26 的习题 4 和 5).

c) 假定 L 是 K 的有限可分扩张 (§26 的习题 4; 我们提醒这个条件在特征为 0 的情形总是满足的). 设 $(u_i)_{1\leqslant i\leqslant n}$ 是由在 A 上整元素组成的 L 的一个基 (将证明这样的基存在), 而 $(v_i)_{1\leqslant i\leqslant n}$ 是补基 (§26 的习题 4). 设 B 是 A 在 L 内的整闭包. 证明所有的 $x \in \mathrm{B}$ 关于基 $(v_i)_{1\leqslant i\leqslant n}$ 的分量在 A 内. 由此推出如果 A 是 Noether 的则 A-模 B 是有限生成的, 并且如果 A 是主理想整环, 则 A-模 B 同构于 A^n.

¶¶48. 设 L 是一个代数数域 (§26 的习题 4), 而 B 是在 $\mathbf{Z}$ 上的整元 $x \in \mathrm{L}$ 的环 (习惯上称 B 为 L 的**整元环**).

a) 证明 B 的加法群具有 n 个元素组成的一个基, 这里 $n = [\mathrm{L} : \mathbf{Q}]$ (即称在 $\mathbf{Q}$ 上的 L 的一个基, 它同时是 $\mathbf{Z}$-模 B 的一个基).

b) 证明对于 B 的所有理想 I, 关系 $\mathrm{I} \neq \{0\}$ 蕴含 $\mathrm{I} \cap \mathbf{Z} \neq \{0\}$ (取一个非零 $x \in \mathrm{I}$, 并且考察它在 $\mathbf{Q}$ 上的极小多项式的第一个非零系数). 由此推出商环 B/I 对于 B 的所有非零理想 I 是有限的.

c) 证明 B 的所有非零素理想 $\mathfrak{p}$ 是极大的, $\mathrm{B}/\mathfrak{p}$ 是一个有限域, 并且 $\mathfrak{p} \cap \mathbf{Z} = p\mathbf{Z}$, 这里 p 是域 $\mathrm{B}/\mathfrak{p}$ 的特征.

[这些经典的结果属于 Dedekind, 而习题 47 的结果是它的容易得到的推广. 域 $\mathrm{B}/\mathfrak{p}$ 有限这个事实部分地解释了有限域研究的重要性, 而且所有的有限域都可以用这种方法得到. 在 Dedekind 之前 Gauss 和 Kummer 曾经谋求得到的 Dedekind 的基本结果之一是代数数的一个域的整元环是一个 Dedekind 整环 (名称由此得来), 换句话说, 环 B 的所有非零理想可以写成素理想的乘积, 并且不计素理想次序是唯一的. 习题 49 和 50 的目的就是给出这个事实的证明. 仅用到下列事实: B 是整闭的 (根据习题 46 的 e) 是显然的) Noether 环 (根据习题 48 问题 a) 是显然的), 其所有的非零素理想是极大的.]

¶¶¶49. 设 A 是一个交换整环, 而 K 是它的分式域. 以下利用在 §10, 习题 14 定义的 A 的分式理想的概念.

a) 称 A 的一个分式理想 I 是**除子的**, 如果它是包含它的主分式理想 (即其形式为 Ax, 这里 $x \in \mathrm{K}, x \neq 0$) 的交集. 证明如果 I 和 J 是除子的, 则 (I:J) 同样是除子的.

证明如果 A 是 Noether 环, 则所有 $x \in (\mathrm{I} : \mathrm{I})$ 在 A 上是整元, 并且由此推出如果 A 是 Noether 的, 并且是整闭的, 则有 $(\mathrm{I} : \mathrm{I}) = \mathrm{A}$.

b) 假定 A 是 Noether 的. 证明 A 至少有一个非零素理想是除子的 (考虑满足条件 $\mathrm{I} \subset \mathrm{A}, \mathrm{I} \neq \mathrm{A}$ 的除子理想 I 的集合, 并且取其中的极大元素).

c) 此后假定 A 是局部的[(*)], Noether 的和整闭的, 并且 A 的仅有的非零素理想是 A 的唯一

[(*)] 称任意交换环 A 为局部环, 如果 A 的非可逆元的集合是 A 的一个理想 $\mathfrak{p}$. 那么这个理想就是 A 的唯一的极大理想, 并且如果 A 是整环, 则 A 的分式域的子环 $\mathrm{A}_{\mathfrak{p}}$ 等于 A 自身. 反之, 如果 A 是一个整环, 并且 $\mathfrak{p}$ 是 A 的一个素理想, 则环 $\mathrm{A}_{\mathfrak{p}}$ 是一个局部环. 局部环主要用于研究代数流形在给定点的 “邻近处” 的性质, 这就解释了表示它的术语的来源.

的极大理想 $\mathfrak{p}$. 一个离散的赋值环 (§8, 习题 6) 满足这些条件, 我们打算证明它的逆命题.

证明 $\mathfrak{p}$ 是除子的, 并且由此得到

$$(A:\mathfrak{p}) \neq A.$$

证明对于所有的 $x \in (A:\mathfrak{p})$ 有

$$x\mathfrak{p} = \mathfrak{p} \quad 或 \quad x\mathfrak{p} = A;$$

由此推出

$$(A:\mathfrak{p}) \cdot \mathfrak{p} = A,$$

从而理想 $\mathfrak{p}$ 是可逆的.

d) 证明 $\mathfrak{p} \neq \mathfrak{p}^2$, 并且对于不属于 $\mathfrak{p}^2$ 的所有 $x \in \mathfrak{p}$ 有 $\mathfrak{p} = Ax$.

e) 证明对于 A 的所有非零理想 a 存在一个整数 n, 使得 a 含于 $\mathfrak{p}^n$ 内, 但是不含于 $\mathfrak{p}^{n+1}$ 内. 利用 $\mathfrak{p}$ 是可逆的这个事实, 证明 $\mathfrak{a} = \mathfrak{p}^n$, 并且由此推出 A 是主理想整环.

f) 证明 A 是离散赋值环.

¶¶¶50. 设 A 是一个交换整环, K 是它的分式域. 假定 A 是 Noether 的和整闭的, 并且 A 的所有非零素理想是极大的. 我们打算证明 A 是一个 Dedekind 整环, 即 A 的所有分式理想是可逆的.

a) 设 $\mathfrak{p}$ 是 A 的一个非零素理想, 借助上一个习题, 证明局部环 $A_\mathfrak{p}$ 是 K 的一个离散赋值环. 设 $v_\mathfrak{p}$ 是这样选取的赋值, 使得

$$v_\mathfrak{p}(K^*) = \mathbf{Z}.$$

证明 A 的元素由下列事实刻画: 对于 A 的所有非零素理想 $\mathfrak{p}$ 都有

$$v_\mathfrak{p}(x) \geqslant 0.$$

b) 证明: 对于所有非零的 $x \in K$, 使得 $v_\mathfrak{p}(x) \neq 0$ 的 $\mathfrak{p}$ 的数目是有限的 (归结为 $x \in A$ 的情形, 并且应用 §18 的习题 6 到 A 的理想 Ax). 证明给定了 A 的有限个两两不同的素理想 $\mathfrak{p}, \mathfrak{p}_1, \cdots, \mathfrak{p}_r$ 和整数 $n_1, \cdots, n_r$, 则存在一个 $x \in A$, 满足关系

$$v_\mathfrak{p}(x) = n_i, \quad v_{\mathfrak{p}_i}(x) \geqslant 0 \quad 对于\ 1 \leqslant i \leqslant r.$$

c) 证明对于 A 的所有非零素理想 $\mathfrak{p}$, 存在一个 $x \in K$, 使得

$$v_\mathfrak{p}(x) = -1, \quad v_\mathfrak{q}(x) \geqslant 0 \quad 对于所有\ \mathfrak{q} \neq \mathfrak{p}$$

(选择一个满足 $v_\mathfrak{p}(x_0) = -1$ 的 x_0, 令 $x = x_0 y$, 借助上一个问题确定 y).

d) 证明 A 的所有非零素理想是可逆的, 因此 A 是一个 Dedekind 整环 (参见 §18, 习题 7).

e) 设 A 是一个 Dedekind 整环, L 是 A 的一个分式域的有限次的可分的扩张, 而 B 是 A 在 L 内的整闭包. 证明 B 是一个 Dedekind 整环. [这个方法应用到 $A = k[K]$, 其中的 k 是一个交换域, 就引导到 Dedekind 整环的例子, 它们跟代数整元的理论中的 Dedekind 整环迥然不同.]

f) 证明环 $\mathbf{Z}[\sqrt{-5}]$ 是一个 Dedekind 整环, 再证明由 3 和 $1+\sqrt{-5}$ 在这个环内生成的理想是素的, 并且不是主理想.

g) 设 k 是一个交换的代数闭域. 令 $A = k[X]$ (这里 X 是 k 上的一个未定元), $K = k(X)$, 而

$$L = K[\sqrt{X^3 + pX + q}],$$

其中的 p 和 q 是 k 的元素 (于是 L 是 K 的一个二次扩张, 它对应于方程为

$$y^2 = x^3 + px + q$$

的 "三次曲线"). 求 A 在 L 内的整闭包. 给定一个 $c \in k$, 设 $\mathfrak{p}$ 是由满足条件 $f(c) = 0$ 的 $f \in \mathrm{A}$ 组成的 A 的素的 (并且是极大的) 理想. 在什么情形由 $\mathfrak{p}$ 在 B 内生成的理想 $\mathfrak{p}\mathrm{B}$ 是素理想? 如果它不是素理想, 如何把它分解成素理想的乘积?

¶¶¶ 51. (用到习题 41 的 Hilbert 零点定理的另一个证明.) 设 K 是一个无限交换域, 而

$$\mathrm{L} = \mathrm{K}[x_1, \cdots, x_n]$$

是包含 K 并且由 K 和有限个元素 $x_1, \cdots, x_n$ 生成的交换整环.

a) 假定存在 x_i 之间的一个代数关系

$$f(x_1, \cdots, x_n) = 0,$$

其中 f 是一个 r 次的系数在 K 内的 n 个未定元的非零多项式. 设 f_r 是 f 的总次数为 r 的齐次部分. 给定 K 上的未定元 $\mathrm{Z}_1, \cdots, \mathrm{Z}_{n-1}, \mathrm{Y}$ 和 K 的元素 $c_1, \cdots, c_{n-1}$, 令

$$f(\mathrm{Z}_1 + c_1\mathrm{Y}, \cdots, \mathrm{Z}_{n-1} + c_{n-1}\mathrm{Y}, \mathrm{Y}) = \sum_{0 \leqslant k \leqslant r} p_k(\mathrm{Z}_1, \cdots, \mathrm{Z}_{n-1})\mathrm{Y}^k.$$

证明多项式 p_r 由

$$p_r(\mathrm{Z}_1, \cdots, \mathrm{Z}_{n-1}) = f_r(c_1, \cdots, c_{n-1}, 1)$$

给定, 并且因此是常量. 证明存在 $c_1, \cdots, c_{n-1} \in \mathrm{K}$, 使得

$$f_r(c_1, \cdots, c_{n-1}, 1) \neq 0$$

(利用 f_r 的齐次性和 §28 的定理 1). 这样就选定了 $c_i \in \mathrm{K}$, 令

$$z_i = x_i + c_i x_n \quad (1 \leqslant i \leqslant n-1);$$

证明 $\mathrm{L} = \mathrm{K}[z_1, \cdots, z_{n-1}, x_n]$, 并且 x_n 在 L 的子环 $\mathrm{K}[z_1, \cdots, z_{n-1}]$ 上是整元.

b) 由此推出下列结果 (Emmy Noether 的 "规范化引理"; 如果 K 是有限的它仍然成立, 但证明要困难得多): 如果 $\mathrm{L} = \mathrm{K}[x_1, \cdots, x_n]$ 是域 K 上的有限生成的整环, 并且各个 $x \in \mathrm{L}$ 不全是在 K 上代数的, 则存在具有下列性质的 L 的 $d \leqslant n$ 个元素 $z_1, \cdots, z_d$: (i) 各个 z_i 是系数在 K 内的 x_i 的线性组合, (ii) $z_1, \cdots, z_d$ 在 K 上是代数无关的, (iii) L 的每一个元素在 L 的子环 $\mathrm{K}[z_1, \cdots, z_d]$ 上是整元.

c) 证明: 如果 $x \in \mathrm{L}$ 不是在 K 上代数的, 则环 L 不可能是一个域 (写出 z_d 在 L 内的逆元在子环 $\mathrm{K}[z_1, \cdots, z_d]$ 上是整元, 并且由此推出矛盾). 换句话说, 如果一个在交换域 K 上的有限生成的环 L 是一个域, 那么 L 是 K 的 (必定有限次的) 代数扩张, 如果 K 是代数闭的, 则特别有 $\mathrm{L} = \mathrm{K}$ (由此得到零点定理: §33, 习题 33, e)).

§35 矩阵的典范形式

本节的目的是证明关于矩阵典范形式的 Jordan 定理, 后面会陈述这个定理, 对于不是可对角化的矩阵, 它代替了化成对角形的理论. 虽然这个定理在数学的某些部分 (尤其是线性常微分方程理论) 起着重大作用, 对于初学者来说这方面的知识却不是必需的. 读者在初读时可以忽略本节, 或者仅当作习题来学习.

1. Hamilton-Cayley 定理

设 K 是一个交换环, 而 f 是系数在 K 内的一个未定元的多项式, 即

$$f(\mathrm{X}) = a_0 + a_1\mathrm{X} + \cdots + a_r\mathrm{X}^r.$$

给定一个包含 K 的环 L, 使得 K 是在 L 的中心 (即对于 $a \in \mathrm{K}, u \in \mathrm{L}$ 有 $au = ua$), 还给定一个元素 $u \in \mathrm{L}$, 令

$$f(u) = a_0 + a_1u + \cdots + a_ru^r;$$

如果 L 是交换的, 我们重新回到 §28 第 1 小节的定义 (在一般情形, 可以归结为一个交换环, 只需用系数在 K 内的 u 的幂的线性组合组成的子环 $\mathrm{K}[u]$ 代替 L).

在特殊情形, 如果取 L 为元素在 K 内的 n 阶方阵的环 $\mathrm{M}_n(\mathrm{K})$, 就会发现对于所有多项式 $f \in \mathrm{K}[\mathrm{X}]$ 和所有元素在 K 内的 (甚至在 K 的一个交换的扩张环内的) 方阵 U 可以定义 $f(\mathrm{U})$.

做了这些铺垫, 就可以陈述

定理 1 (Hamilton-Cayley) *设 U 是元素在一个交换环 K 内的方阵, 而*

$$p_{\mathrm{U}}(\mathrm{X}) = \det(\mathrm{U} - \mathrm{X} \cdot 1)$$

是它的特征多项式. 则有

$$p_{\mathrm{U}}(\mathrm{U}) = 0.$$

如果 U 是 n 阶的, 像在 §34 第 3 小节那样, 令

$$(-1)^n p_{\mathrm{U}}(\mathrm{X}) = \mathrm{X}^n - \tau_1(\mathrm{U})\mathrm{X}^{n-1} + \cdots + (-1)^n\tau_n(\mathrm{U}).$$

系数 $\tau_i(\mathrm{U})$ 是带有理整数系数的 U 的元素的多项式, 多项式的系数不依赖 U, 也不依赖基础环 K: 例如用来通过 U 的元素 α_{ij} 计算 $\tau_1(\mathrm{U})$ 的公式

$$\tau_1(\mathrm{U}) = \alpha_{11} + \cdots + \alpha_{nn}$$

对于所有的 n 阶矩阵 U 和所有交换环是一样的. 由于

$$(-1)^n p_{\mathrm{U}}(\mathrm{U}) = \mathrm{U}^n - \tau_1(\mathrm{U})\mathrm{U}^{n-1} + \cdots + (-1)^n\tau_n(\mathrm{U}) \cdot 1_n,$$

我们发现存在含 n^2 个未定元 $X_{ij}(1 \leqslant i, j \leqslant n)$ 的带有理整数系数的多项式 $f_{ij}(1 \leqslant i, j \leqslant n)$, 使得对于所有交换环 K 和所有系数在 K 内的 n 阶矩阵

$$U = (\alpha_{ij})_{1 \leqslant i,j \leqslant n},$$

矩阵 $p_U(U)$ 的元素由把 U 的元素 α_{ij} 代入多项式 f_{ij} 的未定元而得到. 于是为了证明 $p_U(U) = 0$, 只需证明多项式 f_{ij} 是 0, 即它们的系数是 0, 而为此 (§28, 定理 1) 只需证明每当用任意有理整数代入 f_{ij} 都得到 0. 而根据 f_{ij} 的构造本身这意味着对于元素在 **Z** 内的所有方阵 $p_U(U) = 0$.

如此看来, 为了证明定理 1 对于所有交换环 K 成立, 只需证明它对于环 **Z** 成立. 为此我们将对于任意一个代数闭域证明定理, 因此定理将对于 **C** 成立, 从而对于 **C** 的子环 **Z** 成立, 故在所有情形下成立!

为此考虑 K^n 的关于典范基以 U 为矩阵的自同态. 根据 §34 定理 3, 存在 K^n 的一个基 $(x_i)_{1 \leqslant i \leqslant n}$, 使得

$$u(x_i) = \rho_{1i}x_1 + \cdots + \rho_{ii}x_i \quad (1 \leqslant i \leqslant n). \tag{1}$$

由于三角矩阵的行列式是对角线元素的乘积, 故有

$$p_U(X) = p_u(X) = (\rho_{11} - X) \cdots (\rho_{nn} - X),$$

由于 $p_U(U)$ 显然是自同态

$$p_u(u) = (\rho_{11} - u) \cdots (\rho_{nn} - u)$$

(关于 K^n 的典范基) 的矩阵, 故一切都归结为指出 $p_u(u)$ 是 0. 为此, 令

$$u_i = \rho_{ii} - u;$$

从 (1) 推出

$$u_j(x_i) = (\rho_{ii} - \rho_{jj})x_i + y_{ij},$$

其中的 y_{ij} 属于 K^n 的由向量 $x_1, \cdots, x_{n-1}$ 生成的向量子空间 F_{i-1}, 并且从这些公式立刻得到对于所有的 i

$$u_i(F_i) \subset F_{i-1}.$$

为了由此推出自同态

$$v = p_u(u) = u_1 \circ \cdots \circ u_n$$

是零, 只需注意到

$$v(K^n) = v(F_n) = u_1 \circ \cdots \circ u_{n-1} \circ u_n(F_n) \subset u_1 \circ \cdots \circ u_{n-1}(F_{n-1}) \subset \cdots \subset u_1(F_1),$$

由于显然有 $u_1(F_1) = \{0\}$, 定理证毕.

例 1 如果 U 是其元素在一个任意交换环内的二阶方阵, 则有

$$\mathrm{U}^2 - \mathrm{Tr}(\mathrm{U}) \cdot \mathrm{U} + \det(\mathrm{U}) \cdot 1_2 = 0.$$

建议读者通过直接计算验证这个结果.

注 1 由于 $p_{\mathrm{U}}(\mathrm{X}) = \det(\mathrm{U} - \mathrm{X} \cdot 1_n)$, 初学且机敏的读者无疑会有通过写出 

$$p_{\mathrm{U}}(\mathrm{U}) = \det(\mathrm{U} - \mathrm{U} \cdot 1_n) = \det(\mathrm{U} - \mathrm{U}) = \det(0) = 0$$

证明定理的想法. 这个美妙的证明可惜建立在错误的基础上, 错误在于: 当在 $\mathrm{U} - \mathrm{X} \cdot 1_n$ 中把未定元 X 用矩阵 U 代入时得到矩阵 $\mathrm{U} - \mathrm{U} \cdot 1_n = 0$, 其实并非如此; 事实上 $\mathrm{U} - \mathrm{X} \cdot 1_n$ 由 U 的对角线元素减去 X 而得到, 如果把 U 代入所得到的结果中的 X, 就会发现

$$\begin{pmatrix} \alpha_{11} - \mathrm{U} & \alpha_{12} & \cdots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} - \mathrm{U} \end{pmatrix},$$

其元素在由 K 和 U 所生成的 $\mathrm{M}_n(\mathrm{K})$ 的交换子环内, 而这个矩阵显然一般不是 0! Hamilton-Kayley 定理仅断定这个矩阵的行列式 (K[U] 的元素) 是零.

这个注指出有时人们会被自己使用的记号引入歧途.

2. 幂零自同态分解

Jordan 定理证明的第一步在于确立下列结果:

定理 2 设 u 是交换域 K 上的有限维向量空间 E 的一个自同态, 并且假定 u 的特征多项式 p_u 的所有的根都在 K 内. 令

$$p_u(\mathrm{X}) = (\lambda_1 - \mathrm{X})^{r_1} \cdots (\lambda_q - \mathrm{X})^{r_q},$$

其中的 $\lambda_1, \cdots, \lambda_q \in \mathrm{K}$ 是 p_u 的所有不同的根, 而 $r_1, \cdots, r_q$ 是它们的重数. 再令

$$\mathrm{E}_i = \mathrm{Ker}[(u - \lambda_i)^{r_i}] \quad 对于\ 1 \leqslant i \leqslant q.$$

那么 E 是子空间 E_i 的直和, 并且有

$$\dim(\mathrm{E}_i) = r_i \quad 对于\ 1 \leqslant i \leqslant q.$$

为了简化记号, 令

$$p(\mathrm{X}) = p_u(\mathrm{X}), \quad f_i(\mathrm{X}) = (\lambda_i - \mathrm{X})^{r_i},$$
$$g_i(\mathrm{X}) = f_1(\mathrm{X}) \cdots f_{i-1}(\mathrm{X}) f_{i+1}(\mathrm{X}) \cdots f_q(\mathrm{X}).$$

由于 λ_i 是两两不同的, 故多项式 $g_i(1 \leqslant i \leqslant q)$ 是互素的, 因此存在多项式 $h_i \in \mathrm{K}[\mathrm{X}]$, 使得

$$\sum_{1 \leqslant i \leqslant q} h_i(\mathrm{X}) g_i(\mathrm{X}) = 1.$$

令

$$u_i = f_i(u), \quad v_i = g_i(u), \quad w_i = h_i(u),$$

这样得到的 E 的这 $3q$ 个自同态是两两可交换的 (因为 u 的所有多项式都是可交换的), 我们有

$$w_1 \circ v_1 + \cdots + w_q \circ v_q = j_{\mathrm{E}},$$

这里的 j_{E} 是 E 的恒等自同态. 这就表明有

$$x = w_1(v_1(x)) + \cdots + w_q(v_q(x)) \quad \text{对于所有 } x \in \mathrm{E}. \tag{2}$$

而根据定理 1 对于所有的 i, 我们有

$$0 = p(u) = f_i(u) g_i(u) = u_i \circ v_i,$$

由于 u_i 和 w_i 交换, 更加有 $u_i \circ w_i \circ v_i = 0$. 这就表明对于每个 i 和所有 $x \in \mathrm{E}$ 有

$$w_i(v_i(x)) \in \mathrm{E}_i,$$

因而 (2) 表明有

$$\mathrm{E} = \mathrm{E}_1 + \cdots + \mathrm{E}_q.$$

还要指出这个和是直和, 并且有 $\dim(\mathrm{E}_i) = r_i$. 由于

$$r_1 + \cdots + r_q = d^{\circ}(p_u) = \dim(\mathrm{E}),$$

§19 的定理 13 的推论 3 表明只需建立关系

$$\dim(\mathrm{E}_i) \leqslant r_i. \tag{3}$$

由于 $\mathrm{E}_i = \mathrm{Ker}(u_i)$, 并且 u 和 u_i 是交换的, 显然有 $u(\mathrm{E}_i) \subset \mathrm{E}_i$; 用 u'_i 表示由 u 诱导的 E_i 的自同态, 我们发现 (§34, 注 4) p_u 可以被 $p_{u'_i}$ 整除. 由于 p_u 的所有根都在 K 内, 故 $p_{u'_i}$ 也如此, 因此存在 E_i 的一个基, 使得关于这个基, u'_i 的矩阵是三角的 (§34, 定理 3). 但是根据 E_i 的定义,

$$(u'_i - \lambda_i)^{r_i} = 0,$$

u'_i 关于所说的基的矩阵的对角线元素必然等于 λ_i, 由此即得

$$p_{u'_i}(\mathrm{X}) = (\lambda_i - \mathrm{X})^{\dim(\mathrm{E}_i)}.$$

因为这个多项式整除 p_u, 显然必须有关系 (3) 满足, 这就完成了证明.

标量 $\lambda_1, \cdots, \lambda_q$ 显然是 u 的不同的特征值. 定理 2 表明为了 u 关于 E 的一个基的矩阵尽可能简单, 只需在 E_i 内, 对于由 u 或同样由 $u-\lambda_i$ 诱导的 E_i 的自同态解决同样的问题, 而 $u-\lambda_i$ 是幂零的, 意思是它的某次幂是零.

这样一来, 事情归结为给定向量空间的一个幂零自同态, 选择一个基, 使得自同态关于这个基的矩阵尽可能简单. 在下一小节就做这件事情.

3. 幂零自同态的结构

这就是证明 Jordan 定理的第二步:

定理 3 设 u 是交换域 K 上的维数 $n \geqslant 1$ 的向量空间 E 的一个自同态. 假定存在一个整数 $p \geqslant 0$, 使得

$$u^p = 0$$

(即 u 是幂零的). 则存在 E 的一个基, 使得 u 关于这个基的矩阵有形式

$$\begin{pmatrix} 0 & \nu_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \nu_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \nu_n \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

其中的 ν_i 是 0 或者 1.

显然可以假定 $u \neq 0$, 如果 $u = 0$, 定理是平凡的. 于是存在一个整数 $q \geqslant 1$, 使得

$$u^q \neq 0, \quad u^{q+1} = 0.$$

对于所有整数 $r \geqslant 0$, 令

$$E_r = \operatorname{Ker}(u^r),$$

我们有 $E_0 = \{0\}$ 和 $E_{q+1} = E$.

引理 1 子空间序列

$$\{0\} = E_0 \subset E_1 \subset \cdots \subset E_q \subset E_{q+1} = E$$

是严格递增的, 并且对于所有 $i \geqslant 0$ 有 $u(E_{i+1}) \subset E_i$.

如果一个向量 x 被 u^{i+1} 变为零, 那么显然 $u(x)$ 被 u^i 变为零, 这就证明了引理 1 所宣布的第二个断言. 剩下要做的是证明对于 $0 \leqslant i \leqslant q$, E_{i+1} 严格包含 E_i. 首先对于 $0 \leqslant i \leqslant q$ 关系

$$E_{i+1} \supset E_i$$

是平凡的. 现在假定对于一个满足 $0 \leqslant i \leqslant q$ 的指标 i 有 $\mathrm{E}_{i+1} = \mathrm{E}_i$, 那么对于所有 $x \in \mathrm{E}$, 我们有

$$0 = u^{q+1}(x) = u^{i+1}(u^{q-i}(x)),$$

故有 $u^{q-i}(x) \in \mathrm{E}_{i+1}$; 从 $\mathrm{E}_{i+1} = \mathrm{E}_i$ 得到对于所有 $x \in \mathrm{E}$ 有 $u^{q-i}(x) \in \mathrm{E}_i$, 因此对于所有 $x \in \mathrm{E}$ 有 $u^q(x) = 0$, 这与 q 的定义相矛盾.

引理 2 设 i 是一个满足 $1 \leqslant i \leqslant q$ 的指标, 而 F 是 E 的一个向量子空间, 使得 $\mathrm{F} \cap \mathrm{E}_i = \{0\}$, 则有 $u(\mathrm{F}) \cap \mathrm{E}_{i-1} = \{0\}$, 并且 u 诱导一个从 F 到 $u(\mathrm{F})$ 上的同构.

考虑一个向量 $x \in u(\mathrm{F}) \cap \mathrm{E}_{i-1}$, 那么存在一个 $y \in \mathrm{F}$, 使得 $x = u(y)$, 而由于 $u^{i-1}(x) = u^i(y)$, 我们发现关系 $x \in u(\mathrm{F}) \cap \mathrm{E}_{i-1}$ 蕴含 $y \in \mathrm{F} \cap \mathrm{E}_i$, 故有 $y = 0$, 因此 $x = 0$. 由 u 诱导的从 F 到 $u(\mathrm{F})$ 上的映射显然是线性的和满射的. 它还是单射的, 因为如果 $y \in \mathrm{F}$ 满足 $u(y) = 0$, 则 $y \in \mathrm{F} \cap \mathrm{E}_i = \{0\}$, 即 $y = 0$. 引理得证.

引理 3 存在 E 的向量子空间 $\mathrm{F}_1, \cdots, \mathrm{F}_{q+1}$, 具有下列性质:

a) 对于满足 $1 \leqslant i \leqslant q+1$ 的所有的 i, E_i 是 E_{i-1} 和 F_i 的直和;

b) 对于满足 $2 \leqslant i \leqslant q+1$ 的所有的 i, u 单射地映射 F_i 到 F_{i-1} 内.

首先取 E_q 在 $\mathrm{E}_{q+1} = \mathrm{E}$ 的补空间作为 F_{q+1}, 于是子空间 $u(\mathrm{F}_{q+1})$ 包含于 E_q 内, 并且根据引理 2 与 E_{q-1} 仅相交于 0. 因此存在 E_{q-1} 在 E_q 内的补空间 F_q, 使得 F_q 包含 $u(\mathrm{F}_{q+1})$. 根据引理 1, $u(\mathrm{F}_q)$ 包含于 E_{q-1} 内, 并且根据引理 2, 与 E_{q-2} 仅相交于 0. 于是可以构造 E_{q-2} 在 E_{q-1} 内的一个补空间 F_{q-1}, 使得 F_{q-1} 包含 $u(\mathrm{F}_q)$. 如此构造下去, 显然得到满足条件 a) 的子空间 F_i 的序列, 并且使得 $u(\mathrm{F}_i) \subset \mathrm{F}_{i-1}$. u 单射地映射 F_i 到 F_{i-1} 内这一事实则从引理 2 的第二个断言得到.

现在可以完成定理 3 的证明. 为此, 构造引理 3 的子空间 $\mathrm{F}_i (1 \leqslant i \leqslant q+1)$. 用

$$x_{11}, x_{12}, \cdots, x_{1,r_1}$$

表示 F_{q+1} 的一个基. 由于这些向量是线性无关的, 并且 u 单射地把 F_{q+1} 映射到 F_q 内, 它们在同态 u 下的像是线性无关的, 因而构成 F_q 的基的一部分 (§19, 定理 2). 换句话说存在 F_q 的基, 其形式是

$$x_{21}, x_{22}, \cdots, x_{2r_1}, x_{2,r_1+1}, \cdots, x_{2r_2},$$

其中的

$$u(x_{1j}) = x_{2j} \quad 对于\ 1 \leqslant j \leqslant r_1.$$

从 $x_{2j} (1 \leqslant j \leqslant r_2)$ 出发进行同样的推理, 我们发现存在 F_{q-1} 的一个基,

$$x_{31}, \cdots, x_{3r_3},$$

使得

$$u(x_{2j}) = x_{3j} \quad \text{对于 } 1 \leqslant j \leqslant r_2.$$

如此这般进行下去, 最后得到 $F_1 = E_1$ 的一个基

$$x_{q+1,1}, \cdots, x_{q+1,r_{q+1}},$$

使得有

$$u(x_{q,j}) = x_{q+1,j} \quad \text{对于 } 1 \leqslant j \leqslant r_q.$$

由于 $E_1 = \mathrm{Ker}(u)$, 还有

$$u(x_{q+1,j}) = 0 \quad \text{对于 } 1 \leqslant j \leqslant r_{q+1}.$$

说到此, 由于 E 显然是 $F_1, \cdots, F_{q+1}$ 的直和, 我们发现这样构造的 $r_1 + r_2 + \cdots + r_{q+1}$ 个向量 x_{ij} 组成 E 的一个基. 把这个基写成表的形式就是这样:

$$\begin{array}{lllllllll}
x_{11} & \cdots & x_{1r_1} & & & & & & \\
u(x_{11}) & \cdots & u(x_{1r_1}) & x_{2,r_1+1} & \cdots & x_{2r_2} & & & \\
\cdots\cdots & & & & & & & & \\
u^q(x_{11}) & \cdots & u^q(x_{1r_1}) & u^{q-1}(x_{2,r_1+1}) & \cdots & u^{q-1}(x_{2r_2}) & \cdots & x_{q+1,r_{q+1}}
\end{array}$$

从底部开始, 逐列写出这些向量, 我们找到 E 的一个基 $(x_i)_{1\leqslant i\leqslant n}$, 并且具有性质: 对于每个 i 有

$$u(x_i) = 0 \quad \text{或} \quad u(x_i) = x_{i-1},$$

u 关于这个基的矩阵就具有定理 3 所断言的形式.

4. Jordan 定理

在本节定理之前, 先引进下列定义: 称形式为

$$\begin{pmatrix}
\lambda & 1 & 0 & 0 & \cdots & 0 & 0 \\
0 & \lambda & 1 & 0 & \cdots & 0 & 0 \\
0 & 0 & \lambda & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & \lambda & 1 \\
0 & 0 & 0 & 0 & \cdots & 0 & \lambda
\end{pmatrix}$$

(其中 $\lambda \in K$) 的方阵为元素在一个交换域 K 内的**简化矩阵** (或 **Jordan 矩阵**), 即其对角线元素相等, 对角线上方的紧邻元素为 1, 其余元素全为 0.

例如,

$$(\lambda),\quad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},\quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

分别是一、二、三阶简化矩阵.

定理 4 (Jordan)　设 E 是交换域 K 上的有限维向量空间, 而 u 是 E 的自同态. 则下列性质是等价的:

a) u 的所有特征值在 K 内, 即多项式 p_u 的所有的根在 K 内.

b) 存在 E 的一个基, 关于这个基 u 的矩阵有形式

$$\begin{pmatrix} \mathrm{U}_1 & 0 & 0 & \cdots & 0 \\ 0 & \mathrm{U}_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \mathrm{U}_r \end{pmatrix},$$

其中的每个 U_i 是元素在 K 内的简化矩阵.

假定 u 的所有特征值在 K 内, 首先应用定理 2. 设 u 在 E_i 的限制是 u_i. 对于每个 i, 假定构造了 E_i 的一个基 $(x_{ij})_{1\leqslant j\leqslant n_i}$, 使得 u_i 关于这个基有定理陈述中所指出的形式, 那么合并这样得到的各个 E_i 的基, 就找到 E 的一个基, 使得关于这个基 u 的矩阵具有所要求的形式. 于是事情归结为考察 u_i. 我们有

$$u_i = \lambda_i + v_i,$$

其中 v_i 是 E_i 的幂零自同态, 于是只需指出存在 E_i 的一个基, 使得 v_i 关于这个基的矩阵是简化矩阵的一个分块对角矩阵. 为此对于 v_i 和 E_i 利用定理 3, 并且每当有一串 ν_i 等于 1 就组成一个简化矩阵. 例如

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

写成

$$\begin{pmatrix} \mathrm{U}_1 & & \\ & \mathrm{U}_2 & \\ & & \mathrm{U}_3 \end{pmatrix},$$

其中的

$$\mathrm{U}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathrm{U}_2 = (0), \quad \mathrm{U}_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

显然如此进行就证明了 a) 蕴含 b).

为了证明 b) 蕴含 a), 我们注意到根据 §34 注 4, 性质 b) 蕴含

$$p_u(\mathrm{X}) = p_{\mathrm{U}_1}(\mathrm{X}) \cdots p_{\mathrm{U}_r}(\mathrm{X}),$$

于是为了证明 p_u 的所有根都在 K 内, 只需指出当 U 是简化矩阵时如此, 而像在 §34 第 5 小节所看到的那样这是显然的.

Jordan 定理的假设 a), 同样还有性质 b), 当 K 是代数闭域时总是满足的. 在一般情形, 条件 a) 还意味着 (§34, 定理 3) u 是可三角化的.

不言而喻, Jordan 定理还可以应用到矩阵: 如果 U 是一个其元素在 K 内的 n 阶方阵, 并且 U 的所有特征值都在 K 内, 则存在一个矩阵 $\mathrm{P} \in \mathrm{GL}(n, \mathrm{K})$, 使得

$$\mathrm{PUP}^{-1} = \begin{pmatrix} \mathrm{U}_1 & 0 & 0 & \cdots & 0 \\ 0 & \mathrm{U}_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \mathrm{U}_r \end{pmatrix},$$

其中的 U_i 是简化矩阵.

例 2 设 U 是其元素在一个代数闭域内的四阶方阵. 那么 U 相似于下列矩阵之一:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix},$$

$$\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \mu & 1 \\ 0 & 0 & 0 & \mu \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}.$$

初看似乎还有其他的可能性, 但可以化到前面中的某一个, 比如矩阵

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix} \quad \text{相似于} \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix},$$

通过交换坐标轴就可以验证这个事实.

§35 习题

把下列矩阵化成 Jodan 型 (在每一个情形计算允许化成 Jodan 型的基的变换):

1. $\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$.
2. $\begin{pmatrix} 12 & -6 & -2 \\ 18 & -9 & -2 \\ 18 & -9 & -3 \end{pmatrix}$.
3. $\begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$.
4. $\begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}$.
5. $\begin{pmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix}$.
6. $\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & n & n-1 & \cdots & 2 \\ n & n-1 & n-2 & \cdots & 1 \end{pmatrix}$.

¶7. 证明如果

$$\mathrm{A} = \begin{pmatrix} a & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & a & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & a \end{pmatrix}$$

是一个 n 阶 Jodan 矩阵, 而 $f(\mathrm{X})$ 是一个未定元的多项式, 则有

$$f(\mathrm{A}) = \begin{pmatrix} f(a) & f_1(a) & f_2(a) & \cdots & f_{n-1}(a) \\ 0 & f(a) & f_1(a) & \cdots & f_{n-2}(a) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & f(a) \end{pmatrix},$$

其中我们令

$$f_k(a) = f^{(k)}(a)/k!$$

(这里假定基础域的特征为 0; 基础域的特征非零时将会怎样?)

¶8. 设 A 是其元素在一个交换域 K 内的矩阵. 不利用 Hamilton-Cayley 定理, 证明存在非常量的多项式 $f \in \mathrm{K}[\mathrm{X}]$, 使得 $f(\mathrm{A}) = 0$. 证明 f 是其中的次数最低者的倍式, 如果要求其中的次数最低的多项式的首项系数为 1, 那么它是唯一的. 这个多项式称为**在 K 上 A 的极小多项式**, 它整除 A 的特征多项式, 因此其次数至多等于 A 的阶.

证明如果 A 是上题中的 Jodan 矩阵, 则 A 的极小多项式是

$$(\mathrm{X} - a)^n.$$

证明如果

$$\mathrm{A} = \begin{pmatrix} \mathrm{A}' & 0 \\ 0 & \mathrm{A}'' \end{pmatrix},$$

则 A 的极小多项式是 A' 和 A'' 的极小多项式的最小公倍式.

证明两个相似的多项式 A 和 PAP^{-1} 具有相同的极小多项式.

假定 K 是代数闭的, 利用 Jordan 定理从前面的结果推出任意方阵的极小多项式的计算方法. 把这个方法应用于前面的习题 1 至 6 的矩阵.

9. 设 L 是一个交换域, K 是 L 的子域, 而 A 是其元素在 K 内的一个方阵. A 在 K 上的极小多项式是否等于 A 在 L 上的极小多项式?

¶¶10. 设 k 是一个交换域, V 是 k 上的有限维向量空间, 而 u 是 V 的一个自同态. 考虑多项式环 $\mathrm{K} = k[\mathrm{X}]$, §§27, 28 的习题 19 中定义的 K-模 V[X], 以及 §§27, 28 的习题 20 中定义的 K-模 V_u. 给定 $x \in \mathrm{V}$ 和 $f \in \mathrm{K}$, 用

$$f \cdot x = f(u)(x)$$

表示在 K-模 V_u 内 x 乘以 f 的乘积, 此外用 $\bar{u}$ 表示由

$$\bar{u}(m_0 + m_1\mathrm{X} + \cdots) = u(m_0) + u(m_1)\mathrm{X} + \cdots \quad \text{对于任意几乎全部为零的 } m_i \in \mathrm{V},$$

给定的 K-模 V[X] 的自同态.

a) 考虑由

$$\theta(m_0 + m_1\mathrm{X} + m_2\mathrm{X}^2 + \cdots) = m_0 + u(m_1) + u^2(m_2) + \cdots$$

给定的映射

$$\theta : \mathrm{V}[\mathrm{X}] \to \mathrm{V}_u;$$

证明 θ 是 K-模的一个同态, 并且 θ 是满射.

b) 证明 θ 的核等于 V[X] 自同态

$$\bar{u} - \mathrm{X} \cdot j$$

的像 (这里 j 表示 V[X] 到自身的恒等映射, 因此 $\mathrm{X} \cdot j$ 是在这个 $k[\mathrm{X}]$ 内的比例为 X 的位似).

c) 设 $\mathrm{A} = (a_{ij})_{1 \leqslant i,j \leqslant n}$ 是 u 关于在 k 上的 V 的一个基的矩阵, 考虑元素在环 K 内的矩阵

$$\mathrm{A} - \mathrm{X} \cdot 1_n = \begin{pmatrix} a_{11} - \mathrm{X} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} - \mathrm{X} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} - \mathrm{X} \end{pmatrix}.$$

借助 §32 的习题 15 证明存在矩阵 $\mathrm{P}, \mathrm{Q} \in \mathrm{GL}(n, \mathrm{K})$, 使得

$$\mathrm{P}(\mathrm{A} - \mathrm{X} \cdot 1_n)Q = \begin{pmatrix} d_1(\mathrm{X}) & 0 & \cdots & 0 \\ 0 & d_2(\mathrm{X}) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n(\mathrm{X}) \end{pmatrix},$$

其中的 $d_1, \cdots, d_n$ 是非零多项式, 并且每一个整除下一个 (请注意一般说来 P 和 Q 的元素依赖 X). 证明对于所有的 i, 多项式 $d_1 \cdots d_i$ 是 $\mathrm{A} - \mathrm{X} \cdot 1_n$ 的 i 阶子式的最大公因式. 以下假定每一个 d_i 的首项系数等于 1.

d) 借助问题 b) 和 §32 的习题 15, 证明 K-模 V_u 同构于商模 $\mathrm{K}/d_i\mathrm{K}$ 的直积. 假定

$$d_1 = \cdots = d_s = 1,$$

而 d_{s+1} 不是常量. 对于 $s+1 \leqslant i \leqslant n$ 令

$$d_i(\mathrm{X}) = \mathrm{X}^{n_i} - a_{i,n_i-1}\mathrm{X}^{n_i-1} - \cdots - a_{i0}.$$

最后考虑矩阵

$$\mathrm{A}_i = \begin{pmatrix} 0 & 0 & 0 & \cdots & a_{i,0} \\ 1 & 0 & 0 & \cdots & a_{i,1} \\ 0 & 1 & 0 & \cdots & a_{i,2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_{i,n_i-1} \end{pmatrix}.$$

证明存在 k 上的 V 的一个基, 使得 u 关于这个基的矩阵是

$$\begin{pmatrix} \mathrm{A}_{s+1} & 0 & \cdots & 0 \\ 0 & \mathrm{A}_{s+2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \mathrm{A}_n \end{pmatrix}.$$

证明 d_n 是 u 的极小多项式.

e) 称 $d_1(\mathrm{X}), \cdots, d_n(\mathrm{X})$ 是 u (或 u 关于在 k 上的 V 的任意一个基的矩阵 A) 的**相似不变量**. 证明 V 的两个自同态 u 和 v 是相似的 (即存在 V 的一个自同构 w, 使得 $v = w \circ u \circ w^{-1}$), 必须并且只需它们有同样的相似不变量. [注意到如果 u 和 v 是相似的, 而 A 和 B 是它们关于 V 的任意一个基的矩阵, 则矩阵 $\mathrm{A} - \mathrm{X} \cdot 1_n$ 和 $\mathrm{B} - \mathrm{X} \cdot 1_n$ 在环 $\mathrm{K} = k[\mathrm{X}]$ 上是等价的, 并且应用 §32 习题 15 的 e), 或利用上面的问题 c).]

或表示为: 设 A 和 B 是元素在一个任意的交换域 k 内的两个 n 阶方阵, 存在一个矩阵 $\mathrm{U} \in \mathrm{GL}(n, \mathrm{K})$, 使得

$$\mathrm{B} = \mathrm{UAU}^{-1},$$

必须并且只需对于 $1 \leqslant i \leqslant n$, 矩阵 $\mathrm{A} - \mathrm{X} \cdot 1_n$ 的 i 阶子式的最大公因式等于 $\mathrm{B} - \mathrm{X} \cdot 1_n$ 的 i 阶子式的最大公因式.

(直接推论: 所有矩阵 $\mathrm{A} \in \mathrm{M}_n(k)$ 相似于它的转置矩阵 ${}^t\mathrm{A}$.)

¶¶ 11. 通过计算它们的相似不变量证明矩阵

$$\begin{pmatrix} 3 & 2 & -5 \\ 2 & 6 & -10 \\ 1 & 2 & -3 \end{pmatrix} \quad 和 \quad \begin{pmatrix} 6 & 20 & -34 \\ 6 & 32 & -51 \\ 4 & 20 & -32 \end{pmatrix}$$

是相似的. 对于下列矩阵解答同样的问题:

$$\begin{pmatrix} 4 & 10 & -19 & 4 \\ 1 & 6 & -8 & 3 \\ 1 & 4 & -6 & 2 \\ 0 & -1 & 1 & 0 \end{pmatrix} \quad 和 \quad \begin{pmatrix} 41 & -4 & -26 & -7 \\ 14 & -13 & -91 & -18 \\ 40 & -4 & -25 & -8 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

¶¶12. 设 L 是一个交换域, K 是 L 的一个子域, 而 A, B 是元素在 K 内的两个方阵. 假定 A 和 B 作为元素在 L 内的矩阵是相似的, 证明 A 和 B 作为元素在 K 内的矩阵也是相似的 (利用习题 10). 与 §27 的习题 22 的关系如何?

¶¶13. 设 A 是元素在交换域 K 内的一个 n 阶方阵. 假定 A 的特征多项式的所有的根都在 K 内 (即 A 在 K 上是可三角化的). 证明存在 Jordan 矩阵组成的分块对角矩阵, 它的相似不变量与 A 的相同. 由此和习题 10 的问题 e) 推出 Jordan 定理的一个新的证明.

¶14. 设 A 是元素在一个代数闭域 K 内的 n 阶可逆方阵. 证明在 $\mathrm{GL}(n,\mathrm{K})$ 内存在可对角化矩阵 D 和一个幂幺矩阵 U, 使得

$$\mathrm{A}=\mathrm{DU}=\mathrm{UD},$$

并且 D 和 U 是由 A 唯一决定的 (归结为 Jordan 矩阵). 称 D 和 U 为 A 的**半单**和**幂幺成分**. [可以证明, 举例说如果 K 是特征为 0 的或有限的, 但不必是代数闭的, 则在 K 的一个代数闭的扩张内计算出来的 D 和 U 的元素仍然在 K 内. 例如, 如果 $\mathrm{K}=\mathbf{R}$, 显然如果 $\mathrm{D},\mathrm{U}\in\mathrm{M}_n(\mathbf{C})$ 满足条件, 则它们的虚数共轭也满足条件, 由于 D 和 U 的唯一性, 我们发现事实上 $\mathrm{D},\mathrm{U}\in\mathrm{M}_n(\mathbf{R})$. 自然如果 K 不是代数闭的, 那么 D 是半单的, 未必在 K 上是可对角化的.]

¶15. 设 $\mathrm{X}\in\mathrm{M}_n(\mathrm{K})$, 这里 K 是代数闭的. 证明存在一个可对角化矩阵 D 和一个幂零矩阵 N, 使得

$$\mathrm{X}=\mathrm{D}+\mathrm{N},\quad \mathrm{D}\cdot\mathrm{N}=\mathrm{N}\cdot\mathrm{D},$$

并且这些矩阵由这些条件完全确定. 取 $\mathrm{K}=\mathbf{C}$ 和元素为实数的矩阵 X, 证明 D 和 N 的元素也都是实数. (事实上, 跟习题 14 的情形一样, 可以证明, 如果 X 的元素在一个特征为 0 的或有限的子域内, 则 D 和 N 亦如此.)

¶16. 用 E 表示满足关系

$$f(n+p)=a_{p-1}f(n+p-1)+\cdots+a_0f(n)\quad \text{对所有 } n\in\mathbf{N} \tag{$*$}$$

的所有映射(*)

$$f:\mathbf{N}\to\mathbf{C}$$

的集合, 其中的 $a_0,\cdots,a_{p-1}$ 是给定的复数.

a) 证明存在唯一的一个 $f\in\mathrm{E}$, 对于它, 数 $f(0),\cdots,f(p-1)$ 有给定的值 (没有要求明确地计算 f). 由此推出 E 是从 $\mathbf{N}$ 到 $\mathbf{C}$ 内的所有映射的空间的 p 维向量子空间.

b) 对于 $0\leqslant i\leqslant p-1$, 用 e_i 表示如下定义的 E 的唯一元素:

$$e_i(j)=\begin{cases}0, & \text{如果 } j\neq i;\\ 1, & \text{如果 } j=i,\end{cases}\quad \text{对于 } 0\leqslant j\leqslant p-1.$$

证明 $e_0,\cdots,e_{p-1}$ 组成 E 的一个基.

c) 证明存在唯一的一个 E 的自同态 u, 使得对于所有 $f\in\mathrm{E}$, 函数 $g=u(f)$ 由

$$g(n)=f(n+1)$$

给定. 计算 u 关于 E 的基 $e_0,\cdots,e_{p-1}$ 的矩阵, 并且证明 u 的特征多项式不考虑符号的差别是

$$\mathrm{X}^p-a_{p-1}\mathrm{X}^{p-1}-\cdots-a_0.$$

(*) 这些 “映射” 其实是复数 “序列”, 但是这里把它们看作函数更方便.

d) 证明对于所有整数 $r \geqslant 0$ 和所有 $\lambda \in \mathbf{C}$, E 的子空间

$$\operatorname{Ker}[(u-\lambda)^r]$$

由形如

$$f(n) = \lambda^n g(n)$$

的 $f \in \mathrm{E}$ 组成, 其中 g 是至多 $r-1$ 次的多项式.

e) 设 $\lambda_1, \cdots, \lambda_h$ 是方程

$$\lambda^p - a_{p-1}\lambda^{p-1} - \cdots - a_0 = 0$$

的不同的根, 而 $r_1, \cdots, r_h$ 是它们的重数. 证明: 一个从 $\mathbf{N}$ 到 $\mathbf{C}$ 内映射 f 在 E 内, 即满足递推关系 (*), 必须并且只需存在多项式

$$g_1, \cdots, g_h \in \mathbf{C}[\mathrm{X}],$$

满足条件

$$d^\circ(g_1) < r_1, \cdots, d^\circ(g_h) < r_h,$$

并且有

$$f(n) = g_1(n)\lambda_1^n + \cdots + g_h(n)\lambda_h^n \quad \text{对于所有 } n \in \mathbf{N},$$

如果这些都成立, 则多项式 $g_1, \cdots, g_h$ 完全由 f 确定.

f) 求所有复数序列 $(u_n)_{n \geqslant 0}$, 它们对于所有 $n \geqslant 0$ 有

$$u_{n+5} = u_{n+4} + 5u_{n+3} - u_{n+2} - 8u_{n+1} - 4u_n.$$

¶17. 设 $\mathrm{A} = (a_{ij})_{1 \leqslant i,j \leqslant p}$ 是复元素的 p 阶方阵. 求所有映射

$$f = (f_1, \cdots, f_p) : \mathbf{N} \to \mathbf{C}^p,$$

它们对于 $1 \leqslant i \leqslant p$ 和所有 $n \in \mathbf{N}$ 满足

$$f_i(n+1) = \sum_{j=1}^{p} a_{ij} f_j(n)$$

(把 A 化成 Jordan 标准形). 令关系 (*) 的所有的解对应在 $\mathbf{C}^p$ 取值的函数

$$(f(n), f(n+1), \cdots, f(n+p-1)),$$

利用这里得到的结果重新发现上一个习题的结果.

¶18. 求满足下列关系的从 $\mathbf{N}$ 到 $\mathbf{C}^4$ 内的所有函数:

$$\begin{aligned}
f_1(n+1) &= -5f_1(n) - 3f_2(n) - 2f_3(n) + 4f_4(n),\\
f_2(n+1) &= 2f_1(n) \qquad\qquad\quad + f_3(n) - f_4(n),\\
f_3(n+1) &= 10f_1(n) + 7f_2(n) + 4f_3(n) - 9f_4(n),\\
f_4(n+1) &= 2f_1(n) \qquad\qquad\quad + f_3(n)
\end{aligned}$$

(利用上一个习题).

¶19. 设 K 是一个特征为 0 的一个代数闭域. 在这个习题(*)中, 考虑系数在 K 内的一个未定元的形式幂级数 (§§27, 28, 习题 11).

a) 给定系数在 K 内的一个未定元 T 的形式幂级数

$$x = \sum_{n\in\mathbf{N}} f(n)\mathrm{T}^n/n!$$

(这里 f 是从自然数集合 $\mathbf{N}$ 到 K 内的一个映射), 称形式幂级数

$$x' = \sum_{n\in\mathbf{N}} f(n+1)\mathrm{T}^n/n!$$

为形式幂级数 x 的导元. 证明映射 $x \to x'$ 是环 K[[T]] 的一个导子. 在下面, 我们引进 x 的逐次导元的记号

$$x'' = (x')', x''' = (x'')', \cdots, x^{(r)} = (x^{(r-1)})'.$$

b) 给定常量 $a_0, \cdots, a_{p-1} \in \mathrm{K}$, 证明求满足常系数的齐次线性微分方程

$$x^{(p)} = a_{p-1}x^{(p-1)} + \cdots + a_0 x \tag{**}$$

的形式幂级数的解的问题归结为求解习题 16 的方程 (∗) 的问题.

c) 对于所有 $\lambda \in \mathrm{K}$, 考虑形式幂级数

$$\exp(\lambda\mathrm{T}) = \sum_{n\in\mathbf{N}} \lambda^n \mathrm{T}^n/n!$$

(其定义显然受到经典指数函数的幂级数展开的启发). 给定形式幂级数

$$x = \sum f(n)\mathrm{T}^n/n!$$

证明以下性质是等价的: (i) $f(n) = g(n)\lambda^n$, g 是 $\mathbf{N}$ 上的系数在 K 内的 r 次多项式函数; (ii) 形式幂级数 x 是级数 $\exp(\lambda\mathrm{T})$ 乘以 T 的系数在 K 内的 r 次多项式函数的乘积. (§§27, 28 的习题 8 可能是有用的.)

d) 设 $\lambda_1, \cdots, \lambda_h$ 是方程

$$\lambda^p = a_{p-1}\lambda^{p-1} + \cdots + a_0$$

在 K 中的不同的根, 而 $r_1, \cdots, r_h$ 是它们的重数. 证明微分方程 (∗∗) 的一般解是

$$x = g_1(\mathrm{T})\exp(\lambda_1\mathrm{T}) + \cdots + g_h(\mathrm{T})\exp(\lambda_h\mathrm{T}),$$

其中的 g_i 是系数在 K 内的至多 $r_i - 1$ 次的多项式.

(*) 习题 19 和以下习题的目的是向读者指出存在于矩阵化简理论和微分方程组之间的联系. 当然考虑到其重要性, 这个主题值得更大篇幅的展开 —— 但是这些展开更多地属于分析而非代数. 形式幂级数的引入符合优秀的传统, 因为 Cauchy-Kowalewska 为了建立带解析系数的微分方程组解的存在性所使用的方法就是首先构造 "形式上" 满足给定方程的幂级数 (即像我们在这里所做的在形式幂级数的范围之内满足方程, 而不管它是否收敛), 然后借助它们的系数的放大估计证明这些形式幂级数收敛. 在这里所研究的方程组的情形, 由于所得到的级数的特别简单的形式, 收敛性的证明 (自然是当 $\mathrm{K} = \mathbf{C}$ 时) 是平凡的.

e) 假定 $\mathrm{K}=\mathbf{C}$. 为了从前述结果得到常系数齐次线性微分方程经典理论, 剩下要做的事情是什么?

f) 现在要找满足方程组

$$x_i' = \sum_{j=1}^{p} a_{ij} x_j$$

的 p 个形式幂级数

$$x_i = \sum_{n\in\mathbf{N}} f_i(n)\mathrm{T}^n/n!,$$

方程中的 a_{ij} 是 K 的给定元素. 证明这个问题的求解归结为习题 17 的求解, 并且用微分方程组的语言解释习题 17 的结果.

在下面的习题中, 要求利用习题 19 的 f) 求解给定的微分方程组:

20. $\begin{cases} x' = 5x - 3y + 2z, \\ y' = 6x - 4y + 4z, \\ z' = 4x - 4y + 5z. \end{cases}$

21. $\begin{cases} x' = 7x - 12y + 6z, \\ y' = 10x - 19y + 10z, \\ z' = 12x - 24y + 13z. \end{cases}$

22. $\begin{cases} x' = x - 3y + 3z, \\ y' = -2x - 6y + 13z, \\ z' = -x - 4y + 8z. \end{cases}$

23. $\begin{cases} x' = 3x - y, \\ y' = x + y, \\ z' = 3x + 5z - 3u, \\ u' = 4x - y + 3z - u. \end{cases}$

24. 求微分方程的解

$$x^{(5)} - x^{(4)} - 5x^{(3)} + x'' + 8x' + 4x = 0.$$

25. 利用习题 16 证明等式

$$\sum_{p=1}^{n} p^2 a^p = \frac{a(a+1)}{(1-a)^3} + \frac{(a-7)n - (2a^2 - 5a + 1)n^2}{2(1-a)^3} a^{n+1}$$

(假定 $a \neq 1$).

26. 设 K 是一个交换域, 而 n 是一个正整数. 用 V 表示次数至多是 n 的系数在 K 内的一个未定元的多项式组成的 (K 上的) 向量空间. V 在 K 上的维数是多少? 用 D 表示从 V 到 V 的把每个多项式变换为它的导多项式的映射. 证明 D 是 V 的幂零自同态, 并且求 V 在 K 上的一个基, 使得关于这个基 D 的矩阵有 Jordan 标准形.

§36 Hermit 型

在整个这一节, 用 K 表示一个交换域, 并且假定给定从 K 到 K 内的一个映射, 记作

$$\lambda \to \lambda^*,$$

并且具有下列性质: 这是一个同态, 即有等式

$$(\lambda+\mu)^* = \lambda^* + \mu^*, \quad (\lambda\mu)^* = \lambda^*\mu^*, \quad 1^* = 1, \tag{1}$$

并且还有

$$(\lambda^*)^* = \lambda. \tag{2}$$

一个这样的映射称为 K 的一个**对合**. 这就是域 K 的一个同构, 并且其平方为恒等同构.

最重要的初等例子如下:

例 1 (实正交情形) 取 $K = \mathbf{R}$, 并且对于所有 $\lambda \in K$, $\lambda^* = \lambda$.

例 2 (复正交情形) 取 $K = \mathbf{C}$, 并且对于所有 $\lambda \in K$, $\lambda^* = \lambda$.

例 3 (复 Hermit 情形) 取 $K = \mathbf{C}$, 并且对于所有 $\lambda \in K$, $\lambda^* = \bar{\lambda}$.

但是还有许多其他可能的情形.

例 4 取 $K = \mathbf{Q}[\sqrt{d}]$, 其中 d 是一个非平方的 (§9) 正整数, 对于任意 $a, b \in \mathbf{Q}$, 令

$$(a + b\sqrt{d})^* = a - b\sqrt{d}.$$

1. 半双线性型, Hermit 型

设 L 是域 K 上的向量空间. 称所有映射 $f: L \times L \to K$ 为 L 上的**半双线性型**, 如果它满足条件

(SQ1) 对于所有 $y \in L$, 从 L 到 K 内的映射 $x \to f(x, y)$ 是线性的.

(SQ2) 对于所有 $x \in L$, 从 L 到 K 内的映射 $y \to f(x, y)^*$ 是线性的.

可以用等式代替这些公理, 这就是

$$f(x' + x'', y) = f(x', y) + f(x'', y), \quad f(\lambda x, y) = \lambda f(x, y),$$
$$f(x, y' + y'') = f(x, y') + f(x, y''), \quad f(x, \lambda y) = \lambda^* f(x, y).$$

当在 K 上选择的对合是恒等映射时, 显然就回到了 §21 的双线性型.

称一个半双线性型是 **Hermit 的**, 如果对于任意 $x, y \in L$ 有

$$f(x, y) = f(y, x)^*. \tag{3}$$

一个特殊情形是

$$f(x,x)=f(x,x)^*; \tag{4}$$

在复 Hermit 情形 (例 3), 数 $f(x,x)$ 由于等于它的共轭数, 故总是实数.

如果 K 上的对合是恒等映射, 关系 (3) 写成

$$f(x,y)=f(y,x); \tag{5}$$

这时就称 f 是 L 上的**对称**双线性型.

例 5 当 L 是 K 上的有限维空间时, 容易构造 L 上的所有半双线性型. 为此选择 L 的一个基 $(a_i)_{1\leqslant i\leqslant n}$, 并且设

$$x=\sum\xi_i a_i,\quad y=\sum\eta_i a_i$$

是 L 的两个向量. 根据 (SQ1), 表达式

$$f_y(x)=f(x,y)$$

是 x 的线性型, 故有

$$f(x,y)=f_y(\sum\xi_i a_i)=\sum\xi_i\cdot f_y(a_i)=\sum\xi_i\cdot f(a_i,y).$$

而根据 (SQ2),

$$f_i(y)=f(a_i,y)^*$$

是 y 的线性型, 故得

$$f(a_i,y)^*=f_i(\sum\eta_j a_j)=\sum_j\eta_j f_i(a_j)=\sum_j\eta_j f(a_i,a_j)^*,$$

两端取对合 "*", 并且利用 (1) 和 (2) 即得

$$f(a_i,y)=\sum f(a_i,a_j)\eta_j^*,$$

代入到前面已经得到的 $f(x,y)$ 的表达式显然得到

$$f(x,y)=\sum_{i,j}\alpha_{ij}\xi_i\eta_j^*,\quad 其中\ \alpha_{ij}=f(a_i,a_j). \tag{6}$$

反之, 直接可以验证对于任意 $\alpha_{ij}\in$ K, 由 (6) 给定的函数 f 是 L 上的一个半双线性型.

例 6 假定 f 是例 5 中和 Hermit 半双线性型, 则有

$$\alpha_{ji}=f(a_j,a_i)=f(a_i,a_j)^*=\alpha_{ij}^*.$$

反之, 关系

$$\alpha_{ji} = \alpha_{ij}^* \quad \text{对所有 } i, j \tag{7}$$

蕴含

$$f(x,y) = \sum \alpha_{ij}\xi_i\eta_j^* = \sum \alpha_{ji}^*\eta_j^*\xi_i = \left(\sum \alpha_{ji}\eta_j\xi_i^*\right)^* = f(y,x)^*,$$

因此 (7) 刻画了 Hermit 半双线性型的特征.

当选择 K 的恒等对合时, 关系 (7) 化为

$$\alpha_{ji} = \alpha_{ij}, \tag{8}$$

此式就刻画了 L 上的对称双线性型的特征.

例 7 让我们置身于实正交的情形 (例 3), 并且取 L 为通常空间的起点为原点的向量的向量空间. 标量积

$$f(x,y) = (x|y)$$

是 L 上的对称双线性型.

例 8 在相对论中, 利用由

$$f(x,y) = \xi_1\eta_1 + \xi_2\eta_2 + \xi_3\eta_3 - c\xi_4\eta_4,$$

定义的 $\mathbf{R}^4$ 的对称双线性型, 其中的 c 是光速, 称为 **Lorentz 型**. 这里涉及的当然是实正交情形.

例 9 在复 Hermit 情形, 取在区间 $0 \leqslant t \leqslant 1$ 定义并且连续的所有复值函数 $x(t)$ 的向量空间 L, 那么

$$f(x,y) = \int_0^1 x(t)\overline{y(t)}dt$$

在 L 上定义一个 Hermit 半双线性型 (它在分析中, 尤其在 Fourier 级数中扮演一个重要角色). 还可以取

$$f(x,y) = \int_0^1 \int_0^1 x(t)\overline{y(t)}\mathrm{K}(s,t)dsdt,$$

其中的 $\mathrm{K}(s,t)$ 是在正方形 $0 \leqslant s,t \leqslant 1$ 上定义并且连续的函数, 并且一次性选定. 所得到的表达式是 L 上的一个半双线性型, 并且当且仅当函数 K 满足对于任意 s,t 有

$$\mathrm{K}(t,s) = \overline{\mathrm{K}(s,t)},$$

这个半双线性型是 Hermit 的.

在本节我们要讲述的代数方法对于建立在这个例子中的 Hermit 型的任何非平凡性质都是不够的, 关于这方面的详细研究要求应用由本章讨论所引出的分析方法, 但这个方法要复杂得多, 它是大部分泛函分析 (Hilbert 空间) 内容的起源.

设 L 是 K 上的有限维向量空间, 而 f 是 L 上的一个半双线性型. 给定 L 的一个基 $(a_i)_{1\leqslant i\leqslant n}$, 出现在 (6) 中的标量

$$\alpha_{ij} = f(a_i, a_j)$$

称为 f 关于所考虑的基的**系数**, 而方阵

$$(\alpha_{ij})_{1\leqslant i,j\leqslant n}$$

则称为 f 关于所提及的基的**矩阵**.

我们称元素在 K 内的矩阵

$$\mathrm{A} = (\alpha_{ij})_{1\leqslant i,j\leqslant n}$$

是 **Hermit 的** (关于在 K 上所考虑的对合), 如果其元素满足上面的关系 (7). 当所选择的对合是恒等映射, 即当

$$\alpha_{ij} = \alpha_{ji}$$

时则称 A 是**对称的**.



注 1 在复 Hermit 情形的一个 Hermit 矩阵称为**复 Hermit 矩阵**, 即一个其元素为复数的方阵, 对于任意 i, j 满足

$$\alpha_{ji} = \overline{\alpha_{ij}}.$$

当 $\mathrm{K} = \mathbf{C}$ 时, 对于一个一般的矩阵使用术语 Hermit 而没有预先明确时, 总默认选择 $\mathbf{C}$ 的对合为取复共轭.

2. 非退化型

设 L 是 K 上的向量空间, 而 f 是 L 上的半双线性型. 对于所有 $y \in \mathrm{L}$, 根据 (SQ1), 函数

$$f_y(x) = f(x, y) \tag{9}$$

是 L 上的线性型: 因此, f 定义一个从 L 到 L* 的映射 (§16, 第 1 小节)

$$\hat{f} : y \to f_y. \tag{10}$$

对于任意 $x, y, z \in \mathrm{L}$ 有

$$f_{y+z}(x) = f(x, y+z) = f(x, y) + f(x, z) = f_y(x) + f_z(x),$$

故

$$f_{y+z} = f_y + f_z,$$

或更愿意写成

$$\hat{f}(y+z) = \hat{f}(y) + \hat{f}(z); \tag{11}$$

此外还有

$$f_{\lambda y}(x) = f(x, \lambda y) = \lambda^* \cdot f(x,y) = \lambda^* \cdot f_y(x),$$

因此, 对于任意 $\lambda \in \mathrm{K}$ 和 $y \in \mathrm{L}$ 有

$$\hat{f}(\lambda y) = \lambda^* \hat{f}(y). \tag{12}$$

如果 K 上的对合是恒等的 (双线性型情形), 我们发现 $\hat{f}$ 是从向量空间 L 到向量空间 L^* 内的一个同态. 在一般情形则不再如此, 我们称从 L 到其对偶空间的映射 $\hat{f}$ 是**半线性的**以表示它满足 (11) 和 (12). 在有限维空间, 半线性映射在进行平凡的修改之后, 具有和线性映射同样的性质.

$\hat{f}$ 的核是满足 $\hat{f}(y) = 0$ 的 $y \in \mathrm{L}$ 的集合, 或满足

$$f(x,y) = 0 \quad 对于所有\ x \in \mathrm{L} \tag{13}$$

的 $y \in \mathrm{L}$ 的集合. 如果 (13) 蕴含 $y = 0$, 则我们称 f 是**非退化的**, 即从 L 到其对偶空间 L^* 的映射 $\hat{f}$ 是单射.

定理 1 设 L 是 K 上的有限维向量空间, f 是 L 上的半双线性型, 而 $\mathrm{A} = (\alpha_{ij})$ 是 f 关于 L 的一个基的矩阵. 则以下性质是等价的:

a) 关系

$$f(x,y) = 0 \quad 对于所有\ x \in \mathrm{L}$$

蕴含 $y = 0$.

b) 关系

$$f(x,y) = 0 \quad 对于所有\ y \in \mathrm{L}$$

蕴含 $x = 0$.

c) 矩阵 A 是可逆的, 即有

$$\det(\alpha_{ij}) \neq 0.$$

d) 从 L 到其对偶空间的映射 $\hat{f}$ 是双射, 即对于 L 上的所有线性型 u, 存在唯一的 $y \in \mathrm{L}$, 使得

$$u(x) = f(x,y) \quad 对于所有\ x \in \mathrm{L}.$$

利用例 5 的记号, 对于所有 x 使得 $f(x,y)=0$ 的 y 显然是方程组

$$\sum_j \alpha_{ij}\eta_j^* = 0 \quad (1 \leqslant i \leqslant n) \tag{14}$$

的解, 而对于所有 y 使得 $f(x,y)=0$ 的 x 显然是方程组

$$\sum_i \alpha_{ij}\xi_i = 0 \quad (1 \leqslant j \leqslant n) \tag{15}$$

的解. 在性质 d) 里, 如果令 $u(a_i)=\nu_i$, 事情归结为确定 y, 使得对于所有 i 有 $f(a_i,y)=\nu_i$, 即解方程组

$$\sum_j \alpha_{ij}\eta_j^* = \nu_i \quad (1 \leqslant i \leqslant n). \tag{16}$$

这些交代清楚了, 那么 d) 的意思是对于 (16) 的任意右端, (16) 都有唯一解. d) 与 c) 和 a) (它表示 (14) 仅有平凡解) 的等价性由 §20 的定理 2 得到. 而 b) 的意思是 (15) 只有平凡解, 故矩阵 (α_{ji}) 是可逆的, 而由于这是矩阵 (α_{ij}) 的转置, 我们发现 b) 和 c) 是等价的, 这就完成了证明.



注 2 a) 的意思是 $\hat{f}$ 的核是零元组成的, 从而 a) 和 d) 的等价性还可以从 §19 定理 3 的推论 1 (该结果显然可以推广到半线性映射) 得到.

例 10 Lorentz 型 (例 8) 是非退化的, 因为它关于 $\mathbf{R}^4$ 的典范基的矩阵

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -c \end{pmatrix}$$

是可逆的, 因为 c 不是零.

3. 同态的伴随同态

设 L 和 M 是 K 上的两个有限维向量空间, 而 f 和 g 分别是 L 和 M 上的两个非退化半双线性型.

设 u 是从 L 到 M 内的一个同态, 考虑表达式

$$g(u(x),y) \quad 对于\ x \in \mathrm{L},\ y \in \mathrm{M}.$$

对于给定的 $y \in \mathrm{M}$, 这显然是 $x \in \mathrm{L}$ 的一个线性型, 因此定理 1 的 d) 指出对于每个 $y \in \mathrm{M}$, 存在唯一的一个 $y' \in \mathrm{L}$, 使得

$$g(u(x),y) = f(x,y') \quad 对于所有\ x \in \mathrm{L}.$$

令 $y' = u^*(y)$, 这就定义了一个映射

$$u^* : \mathrm{M} \to \mathrm{L},$$

并且由定义知道有关系

$$f(x, u^*(y)) = g(u(x), y) \quad \text{对于任意 } x \in \mathrm{L},\ y \in \mathrm{M}. \tag{17}$$

映射 u^* 跟 u 一样是线性的. 事实上, 对于任意 $y, z \in \mathrm{M}$ 有

$$\begin{aligned} f(x, u^*(y+z)) &= g(u(x), y+z) = g(u(x), y) + g(u(x), z) \\ &= f(x, u^*(y)) + f(x, u^*(z)) = f(x, u^*(y) + u^*(z)), \end{aligned}$$

故有 $u^*(y+z) = u^*(y) + u^*(z)$. 同样有

$$f(x, u^*(\lambda y)) = g(u(x), \lambda y) = \lambda^* g(u(x), y) = \lambda^* f(x, u^*(y)) = f(x, \lambda u^*(y)),$$

这表明 $u^*(\lambda y) = \lambda u^*(y)$, 并且证明了我们的断言.

由 (17) 定义的同态

$$u^* : \mathrm{M} \to \mathrm{L}$$

称为同态

$$u : \mathrm{L} \to \mathrm{M}$$

的关于型 f 和 g 的伴随同态. 在 $\mathrm{L} = \mathrm{M}, f = g$ 这一特殊情形, 则 u^* 简称为 **u 关于 f 的伴随同态**. 同态的伴随同态的概念对应伴随矩阵的概念. 在前面, 假定 L 具有一个基 $(a_i)_{1 \leqslant i \leqslant p}$, 关于这个基 f 的表达式是

$$f(x, y) = \sum_{1 \leqslant i \leqslant p} \xi_i \eta_i^*,$$

而 M 具有一个基 $(b_j)_{1 \leqslant j \leqslant q}$, 关于这个基 g 的表达式是

$$g(z, t) = \sum_{1 \leqslant j \leqslant q} \zeta_j \tau_j^*,$$

最后设 u 和 u^* 关于所考虑的基的矩阵是

$$\begin{aligned} \mathrm{A} &= (\alpha_{ji})_{1 \leqslant j \leqslant q, 1 \leqslant i \leqslant p}, \\ \mathrm{B} &= (\beta_{ij})_{11 \leqslant i \leqslant p, \leqslant j \leqslant q}. \end{aligned}$$

取向量

$$x = \sum \xi_i a_i, \quad y = \sum \eta_j b_j,$$

则有

$$u(x)=\sum\alpha_{ji}\xi_i b_j,\quad u^*(y)=\sum\beta_{ij}\eta_j a_i,$$

因此有

$$g(u(x),y)=\sum\alpha_{ji}\xi_i\eta_j^*,$$
$$f(x,u^*(y))=\sum\xi_i\beta_{ij}^*\eta_j^*;$$

令所得到的这两个等式右端相等, 这个等式对于任意 x 和 y 成立, 故对于任意 i,j 得关系

$$\beta_{ij}^*=\alpha_{ji},$$

或

$$\beta_{ij}=\alpha_{ji}^*,$$

即 B 由 A 的转置矩阵 ${}^t\mathrm{A}$ 的元素经过 K 的对合变换而得到. 我们称 B 是关于给定的**对合矩阵 A 的伴随矩阵**, 写成

$$\mathrm{B}=\mathrm{A}^*.$$

(每当 K 上所选择的对合是恒等映射, 比如在实正交和复正交情形, 就说转置代替伴随, 并且用通常的记号

$${}^t\mathrm{A}$$

代替 A*.)

例如, 在复 Hermit 情形, **复矩阵 A 的伴随矩阵**就是 A 的复共轭矩阵的转置

$$\mathrm{A}^*={}^t\bar{\mathrm{A}}.$$



注 3　通过 u 的矩阵计算 u^* 的矩阵而不对于 L 和 M 的基做假设 (前面的计算假定了对于 f 和 g 基是后面将看到的所谓 “正交的”) 是可能的. 参见习题 11.

伴随矩阵的计算法则类似于转置矩阵的计算法则, 更明确地说, 就是有关系

$$(\mathrm{A}+\mathrm{B})^*=\mathrm{A}^*+\mathrm{B}^*,\quad(\lambda\mathrm{A})^*=\lambda^*\mathrm{A}^*,$$
$$(\mathrm{AB})^*=\mathrm{B}^*\mathrm{A}^*,\quad 1_n^*=1_n,$$
$$(\mathrm{A}^*)^*=\mathrm{A}.$$

为了得到这些关系, 首先对于转置写出相应的关系 (§16, 定理 4), 再对所得到的关系用 K 的同构

$$\lambda\to\lambda^*$$

进行变换.

还可以对于同态本身验证类似的公式. 例如设 u 和 v 是从 L 到 M 内的两个同态, 则显然对于任意标量 λ 和 μ 有

$$(\lambda u + \mu v)^* = \lambda^* u^* + \mu^* v^*.$$

设 L, M, N 是 K 上的三个有限维向量空间, 它们分别配备了非退化的半双线性型 f, g, h, 如果有同态 $u : \mathrm{L} \to \mathrm{M}$ 和 $v : \mathrm{M} \to \mathrm{N}$, 则同样有公式

$$(v \circ u)^* = u^* \circ v^*.$$

事实上, 令 $v \circ u = w$, 则有

$$f(x, w^*(z)) = h(w(x), z) = h(v(u(x)), z) = g(u(x), v^*(z)) = f(x, u^*(v^*(z))),$$

因此即得想要的关系.

设 u 是从 L 到 M 内的一个同态, "显然的" 公式

$$(u^*)^* = u$$

仅当 f 和 g 是 Hermit 型时才是有效的. 事实上, 令 $u^* = v$, 从 M 到 L 内的这个同态的伴随同态由关系 (17) 给定, 不过应当用 f 代替 g, 用 g 代替 f, u 用 v 代替, 即该式写成

$$g(y, v^*(x)) = f(v(y), x).$$

如果 f 是 Hermit 的, 那么右端根据 (17) 还可以写成

$$f(x, v(y))^* = f(x, u^*(y))^* = g(u(x), y)^*;$$

而如果 g 也是 Hermit 的, 则最后得到

$$g(y, v^*(x)) = g(y, u(x)),$$

由此得到想要的结果, 即如果 $v = u^*$, 则 $v^* = u$.

4. 关于非退化 Hermit 型的正交性

设 f 是 K 上的向量空间 L 上的一个 Hermit 型. 称两个向量 $x, y \in \mathrm{L}$ 关于 f 是**正交的**, 如果

$$f(x, y) = 0.$$

在例 7 的情形下我们重新回到通常的几何的正交概念, 因为 $(x|y)$ 是 x 和 y 的长度以及它们夹角余弦的乘积, 仅当其中一个向量为零或它们之间的夹角为 $\pi/2$ 时才为零.

如果 x 和 y 是正交的, 则有

$$f(x+y,x+y)=f(x,x)+f(y,y),$$

这是因为由于 f 是 Hermit 型, 对任意 x,y, 有

$$\begin{aligned}f(x+y,x+y)&=f(x,x)+f(x,y)+f(y,x)+f(y,y)\\&=f(x,x)+f(x,y)+f(x,y)^*+f(y,y).\end{aligned}$$

在例 7 的情形下, 上面的关系的意思是 $x+y$ 的长度的平方等于 x 和 y 的长度平方的和, 即这归结为 Pythagoras 定理.

现在设 M 是 L 的一个向量子空间, 称正交于所有的 $y\in \mathrm{M}$ 的 $x\in \mathrm{L}$ 的集合为 M 关于 f 的**正交补空间**, 记作

$$\mathrm{M}^{\perp}.$$

根据 (SQ1), 这确实是 L 的一个向量子空间. 在例 7 的情形下, 如果 M 是一条直线 (对应的, 平面), 则 $\mathrm{M}^{\perp}$ 是垂直于 M 的并且过原点的平面 (对应的, 直线).

总有

$$\mathrm{M}\subset(\mathrm{M}^{\perp})^{\perp},\tag{18}$$

这是由于右端由正交于所有 $y\in\mathrm{M}^{\perp}$ 的 $z\in \mathrm{L}$ 组成, 而由于这些 y 都正交于 $x\in\mathrm{M}$, 故显然所有 $x\in\mathrm{M}$ 在关系 (18) 的右端之内.

定理 2　*设 f 是 K 上的一个有限维向量空间 L 上的一个非退化 Hermit 型, 则对于 L 的所有向量子空间 M 有*

$$(\mathrm{M}^{\perp})^{\perp}=\mathrm{M}.$$

只需建立 (18) 的反向包含关系, 即确立所有的 $z\in(\mathrm{M}^{\perp})^{\perp}$ 属于 M. 为此只需 (§19, 定理 3) 指出如果 L 上的一个线性型 u 在 M 上是零, 则有 $u(z)=0$. 而根据定理 1 的 d), 可以对于某个 $y\in\mathrm{L}$ 写出

$$u(x)=f(x,y).$$

u 在 M 上是零这个条件表明 $y\in\mathrm{M}^{\perp}$, 而 $u(z)=0$ 这个事实即 $f(z,y)=0$ 平凡地从 z 正交于 $\mathrm{M}^{\perp}$ 这个假设得到, 这就完成了证明.

定理 3　*设 M 和 N 是 L 的向量子空间. 在定理 2 的假设之下, 我们有关系*

$$(\mathrm{M}+\mathrm{N})^{\perp}=\mathrm{M}^{\perp}\cap\mathrm{N}^{\perp};\quad (\mathrm{M}\cap\mathrm{N})^{\perp}=\mathrm{M}^{\perp}+\mathrm{N}^{\perp}.$$

由于 $\mathrm{M}+\mathrm{N}$ 由 $x+y$ 组成, 其中 $x\in\mathrm{M}$, $y\in\mathrm{N}$, 又由于

$$f(x+y,z)=f(x,z)+f(y,z),$$

那么第一个关系便直接得到. 为了得到第二个, 注意在第一个关系里用 $\mathrm{M}^{\perp}$ 和 $\mathrm{N}^{\perp}$ 分别代换 M 和 N, 则根据定理 2 得

$$(\mathrm{M}^{\perp}+\mathrm{N}^{\perp})^{\perp}=(\mathrm{M}^{\perp})^{\perp}\cap(\mathrm{N}^{\perp})^{\perp}=\mathrm{M}\cap\mathrm{N},$$

重新应用定理 2 就得到想要的另一个关系.

定理 4 设 f 是 K 上的一个有限维向量空间 L 上的一个非退化 Hermit 型, 则对于 L 的所有向量子空间 M 有

$$\dim(\mathrm{M})+\dim(\mathrm{M}^{\perp})=\dim(\mathrm{L}).$$

如果用 M° 表示 M 在 L 的对偶空间 L^{*} 内的零化子, 即 L 上的满足条件

$$u(x)=0 \quad 对于所有\ x\in\mathrm{M} \tag{19}$$

的线性型 u 的集合, 那么有

$$\dim(\mathrm{M})+\dim(\mathrm{M}^{\circ})=\dim(\mathrm{L})$$

(§19, 定理 9). 因此事情归结为 $\mathrm{M}^{\perp}$ 和 M° 有相同的维数. 而给定了 L 上的一个线性型 u, 则存在唯一的一个 $y\in\mathrm{L}$, 使得对于所有 $x\in\mathrm{L}$ 有 $u(x)=f(x,y)$ (定理 1), 而显然关系 (19) 等价于 $y\in\mathrm{M}^{\perp}$, 所以第 2 小节的映射 $\hat{f}$ 映射 $\mathrm{M}^{\perp}$ 到 M° 上, 因此存在一个从 $\mathrm{M}^{\perp}$ 到 M° 上的半线性双射, 这两个向量空间因此有同样的维数, 这就完成了证明.

定理 5 设 f 是 K 上的一个有限维向量空间 L 上的一个 Hermit 型. 给定 L 的一个子空间 M, 则下列性质是等价的:

a) $\mathrm{M}\cap\mathrm{M}^{\perp}=\{0\}$.

b) f 在 M 上的限制在 M 上是非退化的.

c) 空间 L 是子空间 M 和 $\mathrm{M}^{\perp}$ 的直和, 即所有 $x\in\mathrm{L}$ 以唯一的一种方式写成形式

$$x=y+z,\quad 其中\ y\in\mathrm{M},\quad 并且\ z\in\mathrm{M}^{\perp}.$$

如果 f 是非退化的, 前面的条件还等价于下列条件:

d) 空间 L 是子空间 M 和 $\mathrm{M}^{\perp}$ 的和, 即所有 $x\in\mathrm{L}$ 至少以一种方式写成形式

$$x=y+z,\quad 其中\ y\in\mathrm{M},\quad 并且\ z\in\mathrm{M}^{\perp}.$$

为了证明这个定理, 首先注意 $x\in\mathrm{M}\cap\mathrm{M}^{\perp}$ 是满足关系

$$f(x,y)=0 \quad 对于所有\ y\in\mathrm{M}$$

的 $x \in \mathrm{M}$, 于是条件 a) 和 b) 的等价性是显然的. 同样甚至更显然的是条件 c) 即关系

$$\mathrm{L} = \mathrm{M} \oplus \mathrm{M}^{\perp}$$

蕴含 a). 为了完成证明条件 a), b) 和 c) 的等价性, 只需指出 b) 蕴含 c). 由于 c) 是 (§17, 定理 1) a) 和 d) 同时成立, 又由于已经知道 b) 蕴含 a), 故只需证明 b) 蕴含 d).

设 $x \in \mathrm{L}$, 并考虑 M 上的函数

$$u(y) = f(y, x);$$

这显然是 M 上的一个线性型, 并且由于 f 在 M 上的限制在 M 上是非退化的, 定理 1 指出存在一个 $x' \in \mathrm{M}$, 使得

$$u(y) = f(y, x') \quad \text{对于所有 } y \in \mathrm{M}$$

故有

$$f(y, x) = f(y, x'),$$

或写成对于所有 $y \in \mathrm{M}$

$$f(y, x - x') = 0,$$

因此有 $x - x' \in \mathrm{M}^{\perp}$. 由于 $x' \in \mathrm{M}$, 可知 $x \in \mathrm{M} + \mathrm{M}^{\perp}$, 从而 $\mathrm{L} = \mathrm{M} + \mathrm{M}^{\perp}$, 即条件 d). 假定 f 是非退化的, d) 成立, 则有关系 (§19, 定理 13 的推论 2)

$$\dim(\mathrm{M} + \mathrm{M}^{\perp}) = \dim(\mathrm{M}) + \dim(\mathrm{M}^{\perp}) - \dim(\mathrm{M} \cap \mathrm{M}^{\perp});$$

由假设左端等于 $\dim(\mathrm{L})$; 考虑到定理 4, 我们发现 $\dim(\mathrm{M} \cap \mathrm{M}^{\perp}) = 0$. 于是 d) 蕴含 a), 这就结束了证明.

我们注意, 当 f 是非退化的时候, 上述关系写成

$$\dim(\mathrm{M} + \mathrm{M}^{\perp}) = \dim(\mathrm{L}) - \dim(\mathrm{M} \cap \mathrm{M}^{\perp}),$$

于是在这种情形下性质 a), c) 和 d) 的等价性是直接可以得到的. 但是有时候在实际中需要对于退化的 Hermit 型的定理 5, 并且我们为了建立 b) 蕴含 d) 所使用的推理可以推广到 Hilber 空间 (这是某种配备了一个正定的 Hermit 型的无限维向量空间, 它在分析中起着重要的作用).

重新考虑 K 上的有限维向量空间 L 上的一个非退化的 Hermit 型 f. 称 L 的一个向量子空间 M 是**迷向的**, 如果

$$\mathrm{M} \cap \mathrm{M}^{\perp} \neq \{0\};$$

称 M 是**非迷向的**, 如果

$$\mathrm{M} \cap \mathrm{M}^{\perp} = \{0\}.$$

因此定理 5 表明当且仅当 M 是非迷向的有

$$\mathrm{L} = \mathrm{M} \oplus \mathrm{M}^{\perp} \text{ (直和)}.$$

在这种情形, §17 的第 4 小节表明存在唯一的一个向量空间 L 的自同态 p_{M}, 满足下列条件:

$$\begin{aligned} &p_{\mathrm{M}} \circ p_{\mathrm{M}} = p_{\mathrm{M}}, \\ &p_{\mathrm{M}}(\mathrm{L}) = \mathrm{M}, \\ &p_{\mathrm{M}}(\mathrm{M}^{\perp}) = 0. \end{aligned}$$

对于所有 $x \in \mathrm{L}$ 有

$$x = p_{\mathrm{M}}(x) + y, \quad \text{其中 } y \in \mathrm{M}^{\perp},$$

并且这个关系刻画了 p_{M} 的特征 —— 更准确地说, $p_{\mathrm{M}}(x)$ 是使得 $x - p_{\mathrm{M}}(x)$ 正交于 M 的 M 的唯一的元素. 我们称向量 $p_{\mathrm{M}}(x)$ 是 x 在 M 上的**正交投影**, 而同态 p_{M} 是 M 上的**正交投影**.

例 11 设 M 是由一个向量 $a \in \mathrm{L}$ 生成的直线, 显然当且仅当

$$f(a, a) = 0,$$

M 是迷向的. 这时称 a 是对于 f 的**迷向向量**, 而这些向量的集合 (它显然是过原点的直线的一个并集) 称为 f 的**迷向锥**. 比如当 f 是 Lorentz 型时, 迷向锥是使得

$$x^2 + y^2 + z^2 - ct^2 = 0$$

的 $(x, y, z, t) \in \mathbf{R}^4$ 的集合. 反之, 在例 7 的情形 (通常空间的标量积), 唯一的迷向向量是 0, 此外显然如果三个实数 x, y, z 满足关系

$$x^2 + y^2 + z^2 = 0,$$

则有 $x = y = z = 0$. (反之, 这个方程有非平凡的复数解.)

回到一般情形, 设 a 是 f 的一个非迷向向量. 那么所有 $x \in \mathrm{L}$ 在由 a 生成的直线 M 上的正交投影由下式给定:

$$p_{\mathrm{M}}(x) = \frac{f(x, a)}{f(a.a)} a;$$

事实上, 右端的向量属于 M, 并且

$$f(x - p_{\mathrm{M}}(x), a) = f(x, a) - \frac{f(x, a)}{f(a, a)} f(a, a) = 0,$$

这就证明了 $x - p_{\mathrm{M}}(x)$ 正交于 M.

定理 5 的推论　设 f 是 K 上的一个有限维向量空间 L 上的一个 Hermit 型. 以下性质是等价的:

a) 对于 $x \in \mathrm{L}$, 关系 $f(x,x)=0$ 蕴含 $x=0$ (即 f 不具有任何非零迷向向量).

b) 对于 L 的所有向量子空间 M 有

$$\mathrm{L} = \mathrm{M} \oplus \mathrm{M}^{\perp}.$$

显然对于 L 的所有向量子空间 M, 子空间 $\mathrm{M} \cap \mathrm{M}^{\perp}$ 由迷向向量组成, 故性质 a) 蕴含

$$\mathrm{M} \cap \mathrm{M}^{\perp} = \{0\}.$$

因此根据定理 5 此式蕴含 b). 对于 L 的一维子空间 M 写出 $\mathrm{M} \cap \mathrm{M}^{\perp} = \{0\}$ 即可推知 b) 蕴含 a) 这个事实.



注 4　假定 K 是代数闭域, 并且对于所有 $\lambda \in \mathrm{K}$ 有 $\lambda^* = \lambda$. 设 f 是一个 Hermit 型 (考虑到对于 K 的对合的假设, 即对称双线性型), 并且设 L 的维数至少是 2, 则总存在非零迷向向量.

事实上, 由于 $\dim(\mathrm{L}) \geqslant 2$, 在 L 内可以选择两个不成比例的向量 a 和 b. 对于 $\lambda \in \mathrm{K}$ 有

$$f(a\lambda + b, a\lambda + b) = f(a,a)\lambda^2 + 2f(a,b)\lambda + f(b,b).$$

如果 a 是迷向的, 就没有什么要证的了; 如果 a 不是迷向的, 我们发现使得 $a\lambda+b$ 为迷向向量的 $\lambda \in \mathrm{K}$ 是一个二次方程的根. 由于假定 K 是代数闭的, 所考虑的方程至少有一个根, 而对应的迷向向量 $a\lambda + b$ 不是零向量, 因为 a 和 b 不是成比例的.

这个注特别应用到复正交的情形.

5. 正交基

设 L 是 K 上的有限维向量空间, 而 f 是 L 上的 Hermit 型. 称 L 的所有满足条件

$$f(a_i, a_j) = 0 \quad 对于\ i \neq j$$

的基 $(a_i)_{1\leqslant i\leqslant n}$ 为 L 的 (关于 f 的) **正交基**. 这表明 f 关于所提及的基是对角的, 或 f 关于这个基的表达式的形式是

$$f(x,y) = \sum_{1\leqslant i\leqslant n} \alpha_i \xi_i \eta_j^*.$$

例如, $\mathbf{R}^4$ 的典范基关于 Lorentz 型是正交的.

定理 6 设 f 是 K 上的有限维向量空间 L 上的 Hermit 型. 如果 K 的特征异于 2, 则存在 L 的一个关于 f 正交的基.

如果 $f=0$ 则定理是平凡的, 故我们假定 $f\neq 0$. 由于如果 L 是一维的, 则没有什么要证的, 故采用关于 $n=\dim(\mathrm{L})$ 的归纳推理.

假定找到一个关于 f 的非迷向向量 $a_1\in \mathrm{L}$. 那么 (定理 5) L 是由 a_1 生成的直线和垂直于这条直线的超平面 L′ 的直和. 由于 $\dim(\mathrm{L}')=n-1$, 归纳假设表明 L′ 具有关于 f 在 L′ 上的限制 f' 的一个正交基 $(a_2,\cdots,a_n)$. 那么显然 $(a_1,a_2,\cdots,a_n)$ 是 L 的关于 f 的正交基.

为了完成证明, 剩下的任务就是确立以下结果:

引理 设 f 是 K 上的向量空间 L 上的 hermit 型, 并且假定 K 的特征异于 2. 如果 f 不是零, 则在 L 内存在关于 f 的非迷向向量.

换句话说, 如果 $f\neq 0$, 并且对于所有 $x\in \mathrm{L}$ 有 $f(x,x)=0$, 那么 K 有特征 2.

事实上, 假定对于所有 $x\in\mathrm{L}$ 有 $f(x,x)=0$, 则关系

$$f(x+y,x+y)=f(x,x)+f(x,y)+f(x,y)^*+f(y,y)$$

表明对于任意 $x,y\in\mathrm{L}$ 有

$$f(x,y)+f(x,y)^*=0.$$

用 tx 代替 x, 其中 $t\in\mathrm{K}$, 我们发现标量 $u=f(x,y)$ 满足

$$tu+(tu)^*=0 \quad \text{对于所有 } t\in\mathrm{K}.$$

如果 $f\neq 0$, 可以选择 x 和 y, 使得 $u\neq 0$, 然后取 $t=u^{-1}$ 即得

$$0=1+1^*=1+1,$$

这表明 2 是 K 的特征.

¶**注 5** 如果 K 有特征 2, 而 K 的对合是恒等映射, 那么交错型也是对称的, 所有向量 $x\in\mathrm{L}$ 将是迷向的. 于是 K 的特征不等于 2 的假设对于保证前面的结果的有效性是本质的, 当然在 “实际中” 这个假设总是满足的.

推论 1 设 f 是 K 上有限维向量空间 L 上的一个对称双线性型. 假定 K 是代数闭域并且 2 不是其特征 (例如 $\mathrm{K}=\mathbf{C}$). 那么存在 L 的一个基, 使得 f 关于这个基的表达式是

$$f(x,y)=\xi_1\eta_1+\xi_2\eta_2+\cdots+\xi_r\eta_r,$$

其中 r 是不超过 n 的一个正整数. f 是非退化的, 必须并且只需 $r=n$.

事实上, 选择关于 f 正交的基 $(b_i)_{1\leqslant i\leqslant n}$, 可以假定

$$f(b_i,b_i)=\beta_i\neq 0 \quad 对于\ 1\leqslant i\leqslant r,$$
$$f(b_i,b_i)=0 \quad 对于\ r+1\leqslant i\leqslant n.$$

由于 K 是代数闭的, 存在 $\lambda_i\in \mathrm{K}(1\leqslant i\leqslant r)$, 满足

$$\lambda_i^2 f(b_i,b_i)=1 \quad 对于\ 1\leqslant i\leqslant r.$$

由于 f 是双线性型, 这说明向量 $a_i=\lambda_i b_i$ 满足

$$f(a_i,a_i)=1 \quad 对于\ 1\leqslant i\leqslant r.$$

基 $(a_1,\cdots,a_r,b_{r+1},\cdots,b_n)$ 满足所陈述的条件.

如果 $r<n$, 那么 b_n 正交于所有 $b_i(1\leqslant i\leqslant n)$, 从而正交于所有 $x\in \mathrm{L}$, 因此 f 是退化的, 故如果 f 是非退化的, 必有 $r=n$. 反之, 如果 $r=n$, f 关于我们刚构成的基的矩阵是单位矩阵, 故 f 是非退化的.



注 6 更一般的, 型 (20) 是非退化的, 必须并且只需对于 $1\leqslant i\leqslant n, \alpha_i\neq 0$, 因为这正是 f 关于所考虑的基的矩阵是可逆的充分必要条件.

推论 2 设 f 是 n 维实向量空间 L 上的一个对称双线性型. 则存在满足条件 $p+q\leqslant n$ 非负整数 p 和 q 和 L 的一个基, 使得 f 关于这个基的表达式是

$$f(x,y)=\xi_1\eta_1+\cdots+\xi_p\eta_p-\xi_{p+1}\eta_{p+1}-\cdots-\xi_{p+q}\eta_{p+q},$$

并且当且仅当 $p+q=n$, f 是非退化的.

证明与前一个推论的证明类似. 选择关于 f 正交的一个基, 可以假定

$$f(b_i,b_i)>0 \quad 对于\ 1\leqslant i\leqslant p,$$
$$f(b_i,b_i)<0 \quad 对于\ p+1\leqslant i\leqslant p+q,$$
$$f(b_i,b_i)=0 \quad 对于\ p+q+1\leqslant i\leqslant n.$$

令

$$a_i=\begin{cases}\dfrac{b_i}{\sqrt{f(b_i,b_i)}}, & 1\leqslant i\leqslant p,\\ \dfrac{b_i}{\sqrt{-f(b_i,b_i)}}, & p+1\leqslant i\leqslant p+q,\\ b_i, & p+q+1\leqslant i\leqslant p.\end{cases}$$

我们就发现一个基 $(a_1,\cdots,a_n)$, 使得 f 关于这个基的矩阵显然具有所断言的形式. 当且仅当 $n=p+q$, f 非退化的事实由注 6 推出.

推论 3 设 f 是 n 维复向量空间 L 上的一个 Hermit 型. 则存在满足条件 $p+q \leqslant n$ 的非负整数 p 和 q 和 L 的一个基, 使得 f 关于这个基的表达式是

$$f(x,y) = \xi_1\bar{\eta}_1 + \cdots + \xi_p\bar{\eta}_p - \xi_{p+1}\bar{\eta}_{p+1} - \cdots - \xi_{p+q}\bar{\eta}_{p+q},$$

并且当且仅当 $p+q=n$, f 是非退化的.

证明与前一个推论的证明类似, 只需考虑到, 正如在第一小节关系 (4) 所看到的, 对于所有 $x \in \mathrm{L}$, $f(x,x)$ 是实数.

可以证明 (惯性定律) 在这些陈述中提到的非负整数 p 和 q 仅依赖 f, 跟所选取的基无关. 参见习题 24.

6. 规范正交基

设 f 是 K 上有限维向量空间 L 上的一个 Hermit 型. 称 L 的一个基 (a_i) 关于 f 是**规范正交的**, 如果有

$$f(a_i,a_j) = \begin{cases} 0, & i \neq j, \\ 1, & i = j, \end{cases}$$

即 f 关于所提及的基的矩阵是单位矩阵, 或 f 关于这个基的表达式是

$$f(x,y) = \sum_{1 \leqslant i \leqslant n} \xi_i\eta_i^*,$$

其中的 $n = \dim(\mathrm{L})$.

规范正交基的存在性显然要求 f 是非退化的, 而定理 6 的推论 1 指出例如在复正交的情形条件是充分的 (更一般的, 如果 K 是代数封闭的, 2 不是其特征, 并且在 K 上选择的对合是恒等映射).

让我们考虑实正交的或复 Hermit 情形. 如果 f 具有一个规范正交基, 那么对于 $x \in \mathrm{L}$ 有

$$f(x,x) = \sum_{1 \leqslant i \leqslant n} \xi_i\xi_i^* = \sum_{1 \leqslant i \leqslant n} |\xi_i|^2,$$

因此有

$$f(x,x) > 0 \quad \text{对于所有 } x \neq 0 \in \mathrm{L}.$$

称满足这个条件 (在实正交情形或复 Hermit 情形 $f(x,x)$ 总是实数) 的一个型是**正定的**. 反之, 如果 f 是正定的, 推论 2 的证明表明我们有 $p=n$, 因此 L 具有一个规范正交基. 故有

定理 7 设 L 是有限维实 (对应的, 复) 向量空间, 而 f 是 L 上的一个对称双线性型 (对应的, Hermit 半双线性型). L 具有一个规范正交基, 必须并且只需 f 是正定的.

例 12　例 7 的型 $(x|y)$ 显然是正定的. 对于这个型的规范正交基正是三个互相垂直的单位向量, 一个这样的基称为通常空间内的一个**直角坐标系**.



注 7　如果 (a_i) 是关于 f 的规范正交基, 并且设

$$x = \sum \xi_i a_i \in \mathrm{L},$$

则 x 的坐标 ξ_i 由关系

$$\xi_i = f(x, a_i)$$

给定. 事实上,

$$f(x, a_i) = \sum \xi_j f(a_i, a_i) = \xi_i f(a_i, a_i) = \xi_i.$$

注 8　设 L 是 n 维实 (对应的, 复) 向量空间, 而 f 是 L 上的一个对称双线性型 (Hermit 半双线性型). 则存在 L 的一个基, 使得 f 关于这个基的表达式是

$$f(x, y) = \xi_1 \overline{\eta_1}_1 + \cdots + \xi_p \overline{\eta_p} - \xi_{p+1} \overline{\eta_{p+1}} - \cdots - \xi_{p+q} \overline{\eta_{p+q}},$$

其中的整数 $p \geqslant 0, q \geqslant 0$, 满足 $p + q \leqslant n$.

f 是正定的, 必须并且只需 $p = n, q = 0$. 另一个极端情形是 $p = 0, q = n$, 这是**负定型的**情形, 即有

$$f(x, x) < 0 \quad 对于任意\ x \neq 0 \in \mathrm{L}.$$

更一般地说, f 是**半正定的** (**半负定的**), 如果对于所有 x 有 $f(x,x) \geqslant 0$ (对应的, $f(x,x) \leqslant 0$); 这显然表示有 $q = 0$ (对应的, $p = 0$), 所以正定型 (对应的, 负定型) 是非退化的半正定型 (对应的, 半负定型).

一个既非半正定的又非半负定的型称为非半定的; 这意味着有 $p \geqslant 1$ 并且 $q \geqslant 1$, 或存在向量 x 和 y, 使得

$$f(x, x) > 0, \quad f(y, y) < 0,$$

例如 Lorentz 型就是非定的.

7. Hermit 型的自同构

设 L 是 K 上的向量空间, 而 f 是 L 上的一个非退化 Hermit 型. 向量空间 L 的一个自同构 u 如果满足

$$f(u(x), u(y)) = f(x, y) \quad 对于任意\ x, y \in \mathrm{L},$$

则称 u 为 f 的**自同构**.

由于

$$f(u(x), u(y)) = f(x, u^*(u(y))),$$

其中 u^* 为 u 关于 f 的伴随自同构 (第 3 小节), 我们发现 f 的自同构就是 L 的满足关系

$$u^* \circ u = j_{\mathrm{L}}$$

的自同构, 这里 j_{L} 是恒等映射. 由此以及关系

$$(j_{\mathrm{L}})^* = j_{\mathrm{L}}, \quad (v \circ u)^* = u^* \circ v^*, \quad (u^{-1})^* = (u^*)^{-1}$$

立刻推知 f 的自同构组成 L 的自同构群 GL(L) 的一个子群, 记为

$$\mathrm{GL}(f),$$

并称为 f 的**自同构群**.

不论 u 是否是 f 的自同构, 函数

$$g(x, y) = f(u(x), u(y))$$

都是 L 上的 Hermit 型. 为了表示 u 是 f 的自同构, 即 $g = f$, 只需写出 f 和 g 关于 L 的给定的基 $(a_i)_{1 \leqslant i \leqslant n}$ 的系数是相同的 (当然我们假定了 L 是有限维的). 因此 f 的自同构由以下事实刻画其特征: 对于任意指标 i 和 j 有

$$f(u(a_i), u(a_j)) = f(a_i, a_j).$$

在特殊情形, 如果存在 L 的关于 f 的规范正交基, 那么 u 是 f 的一个自同构, 必须并且只需 u 把它变换为 L (关于 f) 的另一个规范正交基.

如果 K 的对合是恒等映射, 并且取 f 为 K^n 上的双线性型

$$f(x, y) = \sum \xi_i \eta_i,$$

则群 $\mathrm{GL}(f)$ 记作

$$\mathrm{O}(n, \mathrm{K}),$$

并且称为**域 K 上的 n 个变量的正交群**. 在这种情形, 如果 A 是 L 的自同构 u (关于 K^n 的规范正交基) 的矩阵, 则正如我们在第 3 小节已经看到的, u^* 的矩阵是 A 的转置 ${}^t\mathrm{A}$, 故 $\mathrm{O}(n, \mathrm{K})$ 是 $\mathrm{GL}(n, \mathrm{K})$ 的由满足

$${}^t\mathrm{A} \cdot \mathrm{A} = 1_n$$

的矩阵 $\mathrm{A} \in \mathrm{M}_n(\mathrm{K})$ 组成的子群. 这样的矩阵称为**正交的**. 我们注意上述关系蕴含

$$1 = \det({}^t\mathrm{A}) \det(\mathrm{A}) = \det(\mathrm{A})^2,$$

因此有

$$\det(\mathrm{A}) = 1 \quad 或 \quad -1.$$

其行列式为 1 的正交矩阵组成 $\mathrm{O}(n, \mathrm{K})$ 的一个子群, 记为

$$\mathrm{O}^+(n, \mathrm{K}) \quad 或 \quad \mathrm{SO}(n, \mathrm{K}),$$

并且称为**域 K 上的 n 个变量的旋转群**.

如果 $\mathrm{K} = \mathbf{R}$ (对应的, $\mathrm{K} = \mathbf{C}$), 就称为**实** (对应的, 复) **正交群**和**实** (对应的, **复**) **旋转群**.

在域 K 和任意对合的情形, 并且取 f 为 K^n 上的 Hermit 型

$$f(x, y) = \sum_{1 \leqslant i \leqslant n} \xi_i \eta_i^*,$$

那么群 $\mathrm{GL}(f)$ 记作

$$\mathrm{U}(n, \mathrm{K}),$$

并且称为关于 K 上所考虑的对合的域 K 上的 n **个变量的酉群**, 这也是满足条件

$$\mathrm{A}^*\mathrm{A} = 1_n$$

的矩阵 $\mathrm{A} \in \mathrm{M}_n(\mathrm{K})$ 的群, 这样的矩阵称为关于所考虑的对合的**酉矩阵**. 在特殊情形, 对于 $\mathrm{K} = \mathbf{C}$ 和对合 $\lambda \to \bar{\lambda}$, 我们得到 n **个变量的复酉群**, 它由**酉复矩阵** A 组成, A 满足条件

$${}^t\bar{\mathrm{A}} \cdot \mathrm{A} = 1_n.$$

例 13　取通常的空间作为 L, 而 f 是例 7 的型 $(x|y)$; 那么对应的旋转群由 (通常意义下的) 绕原点的旋转即点 O 固定的位移组成.

例 14　称 Lorentz 型的自同构群为 **Lorentz 群**, 更确切地说, 这里指的是 Lorentz 型的自同构群的子群, 它由这样的自同构组成, 这些自同构的行列式为 1, 并且把 $t > 0$ 的 $(x, y, z, t) \in \mathbf{R}^4$ 的集合映射到该集合内. 这个群在物理学中起着重要的作用.

8. 正定 Hermit 型的自同构: 化成对角形

我们要证明以下结果:

定理 8　设 L 是有限维复向量空间, f 是 L 上的正定 Hermit 型, 而 u 是 f 的一个自同构. 则存在 u 的特征向量组成关于 f 的 L 的一个规范正交基.

按照一般的方式用 u^* 表示 L 的一个自同构 u 关于 f 的伴随自同构, f 的自同构用关系

$$u^* = u^{-1}$$

刻画其特征, 因此满足条件

$$u^* \circ u = u \circ u^*.$$

如果 L 的一个自同态 u 满足这个关系, 则说它关于 f 是**规范的**. 交代了这些, 那么定理 8 就显然是下列定理的特殊情形了.

定理 9 设 L 是有限维复向量空间, f 是 L 上的正定 Hermit 型, 而 u 是关于 f 规范的 L 的一个自同态. 则存在由 u 的特征向量组成并且关于 f 规范正交的 L 的一个基.

定理的证明建立在几个引理的基础之上.

引理 1 假定 u 是规范的, 则对于 $x \in \mathrm{L}$, 关系 $u(x) = 0$ 蕴含 $u^*(x) = 0$, 反之亦真. 即 $\mathrm{Ker}(u) = \mathrm{Ker}(u^*)$.

首先注意对于任意 $x, y \in \mathrm{L}$ 有

$$f(u(x), u(y)) = f(u^* \circ u(x), y) = f(u \circ u^*(x), y) = f(u^*(x), u^*(y)),$$

因此

$$f(u(x), u(x)) = f(u^*(x), u^*(x)).$$

如果 $u(x) = 0$, 则右端为零, 由于 f 是正定的, 故 $u^*(x) = 0$. 反之类似地得到关系 $u^*(x) = 0$ 蕴含 $u(x) = 0$.

引理 2 假定 u 是规范的, 而 λ 是 u 的一个特征值. 则 $\bar{\lambda}$ 是 u^* 的特征值, 并且对于 $x \in \mathrm{L}$ 关系 $u(x) = \lambda x$ 等价于 $u^*(x) = \bar{\lambda} x$.

令 $v = u - \lambda \cdot j_{\mathrm{L}}$. 由于 λ 和 L 的所有自同态交换, 简单的计算指出 v 像 u 一样是规范的. 由于 $v^* = u^* - \bar{\lambda} \cdot j_{\mathrm{L}}$, 把引理 1 应用到 v 即得引理 2.

下面用 $\lambda_1, \cdots, \lambda_r$ 表示 u 的所有不同的特征值, 而用 $\mathrm{L}_i (1 \leqslant i \leqslant r)$ 表示由 $u(x) = \lambda_i x$ 的解组成的 L 的子空间.

引理 3 假定 u 是规范的, 那么子空间 $\mathrm{L}_1, \cdots, \mathrm{L}_r$ 是两两正交的.

事实上, 设 $x \in \mathrm{L}_i$ 和 $y \in \mathrm{L}_j$, 根据引理 2 有

$$\lambda_i \cdot f(x, y) = f(u(x), y) = f(x, u^*(y)) = f(x, \lambda_j^*(y)) = \lambda_j f(x, y),$$

由于对于 $i \neq j$ 有 $\lambda_i \neq \lambda_j$, 我们看到 $f(x, y) = 0$, 由此即得引理.

引理 4 设 M 是 L 的在 u 和 u^* 下稳定的子空间, 那么 M 关于 f 的正交补空间 $M^\perp$ 在 u 和 u^* 下也是稳定的.

只需指出, 如果 $x \in L$ 具有性质: 对于所有 $y \in M$ 关系 $f(x,y) = 0$ 成立, 则 $u(x)$ 和 $u^*(x)$ 也具有同样的性质. 而我们有 $f(u(x),y) = f(x,u^*(y))$, 又根据假设关系 $y \in M$ 蕴含 $u^*(y) \in M$, 那么对于所有 $y \in M$ 关系 $f(u(x),y) = 0$ 成立是显然的; 同样证明对于所有 $y \in M$ 关系 $f(u^*(x),y) = 0$ 成立.

为了完成定理 9 的证明, 考虑由 u 的特征向量生成的子空间

$$M = L_1 + \cdots + L_r.$$

根据引理 2, 它在 u 和 u^* 下是稳定的, 从而 $M^\perp$ 在 u 和 u^* 下也稳定的. 如果有 $M^\perp \neq \{0\}$, 那么 u 在 $M^\perp$ 至少有一个特征向量, 于是有 $M \cap M^\perp \neq \{0\}$, 这与 f 是正定的相矛盾, 故有 $M^\perp = \{0\}$, 由于 L 是 M 和 $M^\perp$ 的直和 (定理 5 的推论), 故得

$$L = M.$$

于是 L 是 L_i 的和, 根据 §34 的定理 4 甚至是 L_i 的直和.

这个事实建立之后, 根据引理 3, 为了构造 L 的由 u 的特征向量组成的基, 只需在每个 L_i 内选择一个对于 f 在 L_i 的限制的规范正交基, 而这是可能的 (定理 7), 只要注意到正定型在子空间的限制仍然是正定的这个显然的事实. 这就完成了定理 8 和 9 的证明.

当 u 是 f 的自同构时, 引理 2 表明 $u(x) = \lambda x$ 蕴含

$$u^{-1}(x) = \bar{\lambda} x,$$

由此推出 u 的每一个特征值 λ 满足

$$\lambda^{-1} = \bar{\lambda},$$

即它是模等于 1 的复数.

另一个重要的特殊情形是对于 f 的**自伴**同态 u, 即

$$u^* = u.$$

引理 2 指出这时有

$$\lambda = \bar{\lambda},$$

即 u 的所有特征值是实数.

前面这些结果自然可以翻译成矩阵语言. 取 $L = \mathbf{C}^n$ 和

$$f(x,y) = \sum_{1 \leqslant i \leqslant n} \xi_i \overline{\eta_i},$$

根据定理 7 这并不限制一般性. 设 A 是 u 关于典范基的矩阵, 而 U 是从典范基到 u 的特征向量组成的规范正交基的过渡矩阵, u 关于这个基的矩阵是 $U^{-1}AU$ (§15, 定理 2 的推论), 由此得到

$$A = UDU^{-1}.$$

其中的 D 是对角矩阵. 但是 U 使得典范基 (它对于所考虑的型是规范正交的) 过渡到另一个规范正交基, 正如在第 7 小节所看到的, 矩阵 U 是酉矩阵. 故得

定理 9 的推论 设 A 是元素为复数的 n 阶方阵, 满足条件

$$A^*A = AA^*.$$

则存在一个 n 阶复酉矩阵 U 和一个 n 阶对角矩阵 D, 使得

$$A = UDU^{-1}.$$

A 是酉矩阵 (对应的, Hermit 矩阵), 必须并且只需 D 的对角线元素的模是 1 (对应的, 是实数).

关于最后的断言, 我们观察到从 $A = UDU^{-1}$ 得到

$$A^* = (U^{-1})^*D^*U^* = UD^*U^{-1} = U\bar{D}U^{-1},$$

其中的 $\bar{D}$ 是 D 的复共轭矩阵. A 是 Hermit 的, 必须并且只需

$$UDU^{-1} = U\bar{D}U^{-1},$$

即 $D = \bar{D}$, 即 D 是实的. 而 A 是酉矩阵, 必须且只需

$$1_n = A^*A = U\bar{D}U^{-1}UDU^{-1} = U\bar{D}DU^{-1},$$

即 $\bar{D}D = 1_n$, 这就表明 D 的对角线元素的绝对值是 1.

一个矩阵 $A \in M_n(C)$ 如果满足关系 $A^*A = AA^*$ 则称为**规范的**, 因而一个这样的矩阵是可对角化的.

上述推论自然指出一个 Hermit 矩阵 (自然还有实对称矩阵) 的特征值都是实数. 这个结果蕴含下列

定理 10 设 L 是有限维实向量空间, f 是 L 上的正定对称双线性型, 而 u 是对于 f 的 L 上的自伴同态. 那么存在 L 的由 u 的特征向量组成的对于 f 是规范正交的一个基.

我们按照定理 9 的证明方式进行证明. 由于 $u^* = u$, 引理 1 和 2 是平凡的, 而引理 3 和 4 以同样的方式可以证明仍然是成立的. 像上面那样组成由 u 的特征向量生成的 L 的子空间 M, 事情归结为指出 $M = L$, 即 (定理 5 的推论) $M^{\perp} = \{0\}$. 而为

此像上面一样, 只要指出 L 的在 u 下稳定的非零子空间 N 至少有 u 的一个特征向量. 用 f 在 N 的限制代替 f (这个限制仍然是正定的和对称的), 并且用 u 在 N 的限制代替 u (这个限制关于 f 在 N 的限制仍然是自伴的), 最后都归结为指出在所宣布的假设之下, u 在 L 至少有一个特征向量, 即至少有一个实特征向量.

设 $(a_i)_{1\leqslant i\leqslant n}$ 是对于 f 规范正交的 L 的一个基 (定理 7), 根据第 3 小节的计算结果 u 关于这个基的矩阵 A 是实对称的, 于是它的所有特征值是实数 (定理 9 的推论). 由于 u 的特征值也是 A 的特征值, 证明完成.

推论 设 A 是实对称矩阵, 则存在一个实规范正交矩阵 U 和一个实对角矩阵 D, 使得

$$\mathrm{A} = \mathrm{UDU}^{-1}.$$

为了证明这个推论, 考虑 $\mathbf{R}^n$ 上的双线性型

$$f(x, y) = \sum_{1\leqslant i\leqslant n} \xi_i \eta_i$$

和关于规范基的矩阵为 A 的自同态 u. u 对于 f 是自伴的, 故可以取从 $\mathbf{R}^n$ 的典范基到由 u 的特征向量组成的关于 f 规范正交的一个基的过渡矩阵作为 U.



注 9 在所谓专门数学的经典著作里, 三阶实对称矩阵的所有特征值都是实数这个事实, 经常引用下列事实: 其特征值是三次实系数代数方程的根, 并且一个这样的方程 (更一般的 $\mathbf{R}$ 上的奇数次代数方程) 总至少具有一个实根.

这类证明的最大不妥显然是不能推广到四阶矩阵 (甚至认真说来连二阶的都不可能), 因为偶数次实代数方程可能根本没有实根.

一般结果的一个初等证明可以如下得到. 设 (α_{ij}) 是一个实对称矩阵 (或同样的复 Hermit 矩阵), 而 $\lambda \in \mathbf{C}$ 是这个矩阵的一个特征值, 那么齐次线性方程组

$$\sum_j \alpha_{ij}\xi_j = \lambda \xi_i$$

在 $\mathbf{C}^n$ 具有一个非平凡解. 对于这个解有

$$\sum_{i,j} \alpha_{ij}\xi_i\overline{\xi_j} = \lambda \sum_i \xi_i\overline{\xi_i}.$$

而由于 (α_{ij}) 是 Hermit 的左端是实数, 又由于 ξ_i 不全是零,

$$\sum_i \xi_i\overline{\xi_i} = \sum_i |\xi_i|^2$$

是实数并且 > 0, 因此特征值必定是实数.

为了进行对应的 "几何" 推理, 采用定理 9 的记号写出

$$\lambda \cdot f(x, x) = f(\lambda x, x) = f(u(x), x) = f(x, u(x)) = f(x, \lambda x) = \bar{\lambda} \cdot f(x, x),$$

并且注意到由于 f 是正定的, $f(x,x)\neq 0$, 故

$$\lambda=\bar{\lambda}.$$

9. 迷向向量和不定型

在实正交或复 Hermit 情形, 如果向量空间 L 上的一个 Hermit 型是正定的, 或负定的, 即如果对于 $x\in \mathrm{L}$, 表达式 $f(x,x)$ 保持相同的符号, 并且仅当 $x=0$ 时变为零, 则称为**定的**.

一个型是定的显然没有任何非零迷向向量. 事实上, 这个性质刻画了定的型的特征. 换句话说, *f 是定的型, 必须并且只需关系 $f(x,x)=0$ 蕴含 $x=0$.*

指出非定的 (不是说非半定的) 型具有非零迷向向量也是一样的. 由于 f 不是定的, 故存在非零向量 $a,b\in \mathrm{L}$, 使得

$$f(a,a)\geqslant 0,\quad f(b,b)\leqslant 0,$$

而由于如果 a 或 b 自己就是迷向的, 就没有可证的了, 所以可以假定

$$f(a,a)>0,\quad f(b,b)<0.$$

我们要证明存在一个标量 λ, 使得 $x=a\lambda+b$ 是非零迷向向量. x 是非零向量是显然的, 因为如果 $a\lambda+b$ 是零, 那么

$$f(b,b)=f(-\lambda a,-\lambda a)=|\lambda|^2 f(a,a),$$

就与 $f(a,a)$ 有同样的符号, 这违反了假设. 所以事情归结为证明存在 λ, 使得

$$\begin{aligned}0&=f(a\lambda+b,a\lambda+b)\\&=f(a\lambda,a\lambda)+f(a\lambda,b)+f(b,a\lambda)+f(b,b)\\&=\lambda\bar{\lambda}u+v\lambda+\bar{v}\bar{\lambda}+w,\end{aligned}$$

其中我们用了记号

$$u=f(a,a),\quad v=f(a,b),\quad w=f(b,b).$$

显然有 (参照二次三项式化成平方和)

$$\begin{aligned}\lambda\bar{\lambda}u+v\lambda+\bar{v}\bar{\lambda}+w&=u\left[\left(\lambda+\frac{\bar{v}}{u}\right)\left(\bar{\lambda}+\frac{v}{u}\right)-\frac{\bar{v}v-uw}{u^2}\right]\\&=u\left[\left|\lambda+\frac{\bar{v}}{u}\right|^2-\frac{|v|^2-uw}{u^2}\right];\end{aligned}$$

λ 的这个方程在基础域 K (在实正交情形是 $\mathbf{R}$, 在复正交情形是 $\mathbf{C}$) 内具有一个解, 必须并且只需

$$|v|^2-uw\geqslant 0.$$

由于 $f(a,a)>0$, 而 $f(b,b)<0$, 上述条件是满足的, 我们的断言被证明.



注 10 在前面的证明中还可以用 a 和 b 生成的子空间代替 L, 并且利用定理 6 的推论 2 或 3 (沿用这两个引理的记号, 如果不是 $p=n,q=0$ 和 $p=0,q=n$ 这两种情形, 显然存在非零迷向向量, 而这两种情形恰好是定型的情形).

10. Cauchy-Schwarz 不等式

为了结束这一节, 我们要建立一个非常有用的结果, 尽管它是简单的:

定理 11 设 L 是一个实 (对应的, 复) 向量空间, 而 f 是 L 上的一个对称双线性 (对应的, Hermit 半双线性) 型. 假定

$$f(x,x)\geqslant 0 \quad \text{对于所有 } x\in \mathrm{L},$$

则对于任意 $x,y\in \mathrm{L}$ 有

$$|f(x,y)|^2\leqslant f(x,x)f(y,y). \tag{20}$$

事实上, 对于标量 λ 的所有的值, 表达式

$$f(x\lambda+y,x\lambda+y)=u\lambda\bar{\lambda}+v\lambda+\bar{v}\bar{\lambda}+w \tag{21}$$

都是非负的, 其中

$$u=f(x,x),\quad v=f(x,y),\quad w=f(y,y).$$

由此和 u,w 都非负的事实应当推出

$$|v|^2-uw\leqslant 0.$$

如果 $u=0$, 那么显然 (21) 仅当 $v=0$ 时才能恒非负 (如果 $v\neq 0$, 那么可以选择 λ, 使得 $v\lambda$ 取任意值), 在这种情形要建立的不等式是显然的.

如果 $u\neq 0$, 例如对于

$$\lambda=-\bar{v}/u,$$

表达 (21) 是非负的, 对于 λ 的这个选择, (21) 的值是

$$-\frac{|v|^2-uw}{u}$$

(§9 的计算), 由于 $u>0$, 而上式的值非负的事实就证明了定理.

推论 1 假定对于所有的 x 有 $f(x,x)\geqslant 0$. f 是非退化的, 必须并且只需 f 是正定的.

事实上定理 11 表明 $f(x,x)=0$ 蕴含对于所有 $y\in \mathrm{L}$ 有 $f(x,y)=0$, 因此如果 f 是非退化的, 则 $x=0$.

例 15 取 $\mathrm{L} = \mathbf{C}^n$ 和

$$f(x,y) = \sum_{1 \leqslant i \leqslant n} \xi_i \overline{\eta_i}.$$

我们发现对于任意复数成立的不等式

$$|\xi_1\eta_1 + \cdots + \xi_n\eta_n| \leqslant \sqrt{|\xi_1|^2 + \cdots + |\xi_n|^2}\sqrt{|\eta_1|^2 + \cdots + |\eta_n|^2}.$$

例 16 取例 9 的半双线性型

$$f(x,y) = \int_0^1 x(t)\overline{y(t)}.$$

我们发现经常在分析中使用的不等式

$$\left|\int_0^1 x(t)\overline{y(t)}dt\right| \leqslant \sqrt{\int_0^1 |x(t)|^2 dt}\sqrt{\int_0^1 |y(t)|^2 dt}.$$

注 11 在通常空间的标量积 $(x|y)$ 的情形, 定理 11 表明 x 和 y 的标量积的绝对值小于或等于 x 和 y 的长度的乘积. 这个事实的几何解释是显然的, 因为 $(x|y)$ 等于 x 和 y 的长度与 x 和 y 的夹角的余弦的乘积, 而余弦的绝对值总是小于或等于 1.



推论 2 假定对于所有 $x \in \mathrm{L}$, 有 $f(x,x) \geqslant 0$, 令

$$||x|| = \sqrt{f(x,x)},$$

则对于任意 $x, y \in \mathrm{L}$ 有

$$||x+y|| \leqslant ||x|| + ||y||.$$

事实上, 我们有

$$\begin{aligned} ||x+y||^2 = f(x+y,x+y) &= f(x,x) + f(y,y) + \overline{f(x,y)} + f(x,y) \\ &= f(x,x) + f(y,y) + 2 \cdot \mathrm{Re} f(x,y). \end{aligned}$$

由于 Cauchy-Schwaz 不等式还可以写成形式

$$|f(x,y)| \leqslant ||x|| \cdot ||y||,$$

更有

$$\mathrm{Re} f(x,y) \leqslant ||x|| \cdot ||y||,$$

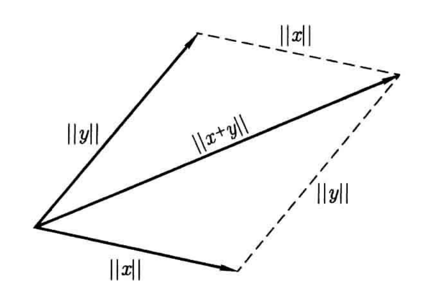
因此有

$$||x+y||^2 \leqslant ||x||^2 + ||y||^2 + 2||x|| \cdot ||y|| = (||x|| + ||y||)^2,$$

由此立刻得到想要的不等式.

在通常空间的标量积 $(x|y)$ 的情形, 显然 $||x||$ 正是向量 x 的长度. 推论 2 证明了下列结果:

推论 3　在一个三角形里, 每一条边小于其他两条边的和.



§36 习题

(习题 1 至 21 是有关于在任意基础域 K 上都成立的性质, 不过读者有时应当排除特征为 2 的域, 我们并没有在每一处明确地指出这一点. 在习题 22 到 51 中我们假定域是 **R** 或 **C**, 而且在每一处都是指明的. 当然在任意基础域 K 上成立的所有结果, 尤其是习题 1, 2, 9 至 21, 当基础域是 **R** 或 **C** 时也是有用的.)

1. 设 V 是交换域 K 上的有限维向量空间, 称 V 上的所有二次齐次多项式函数为 V 上的**二次型** (§28, 习题 17), 即二次型是所有由形如

$$q(x)=\sum a_{ij}x_ix_j$$

的关系给定的从 V 到 K 内的映射, 其中的 a_{ij} 是 K 的给定的元素, 而 $x_i\in \mathrm{K}$ 是向量 $x\in \mathrm{V}$ 关于 V 的一个基的坐标.

a) 证明如果 $f(x,y)$ 是 V 上的一个对称双线性型, 则函数

$$q(x)=f(x,x)$$

是 V 上的一个二次型 (称它是与 f **相伴的**).

b) 假定 K 的特征 $\neq 2$. 证明从 q 可以借助公式

$$f(x,y)=\frac{q(x+y)-q(x-y)}{4}$$

重新构造 f.

c) 反之, 如果 q 是 V 上的一个二次型, 则上面的公式定义 V 上的一个对称双线性型 f, 并且有 $q(x)=f(x,x)$.

d) 一个基关于 f 是正交的, 必须并且只需 q 关于这个基的表达式有形式

$$q(x)=c_1x_1^2+c_2x_2^2+\cdots+c_nx_n^2$$

(这时说 q 化简为一个**平方和**).

[上面的结果指出在特征 $\neq 2$ 的情形对称双线性型的研究等价于二次型的研究. 在下面的习题中我们经常使用二次型的语言.]

2. 设 K 是一个特征 $\neq 2$ 的交换域. 考虑 K^n 上的一个二次型

$$q(x) = \sum_{1 \leqslant i,j \leqslant n} a_{ij} x_i x_j.$$

a) 假定 $a_{11} \neq 0$, 则存在 K^n 上的一个线性型 f_1, 使得

$$q(x) = a_{11} \cdot f_1(x)^2 + q_1(x),$$

其中的二次型 q_1 不再依赖 x_1 (把 $q(x)$ 写成 x_1 的系数依赖 $x_2, \cdots, x_n$ 的二次三项式, 再把这个二次三项式表示成中学生熟知的典范形式).

b) 假定 $a_{11} = 0$, 而 $a_{12} \neq 0$, 作为 K^n 内的新坐标选择线性型

$$y_1 = x_1 + x_2,$$
$$y_2 = x_1 - x_2,$$
$$y_i = x_i \quad (3 \leqslant i \leqslant n),$$

证明

$$q(x) = \sum_{1 \leqslant i,j \leqslant n} b_{ij} y_i y_j,$$

其中的 $b_{11} \neq 0$, 因此在新的坐标系里可以应用问题 a) 的技巧.

c) 由前述内容得到一个实际的方法, 以便把二次型化为平方和 (见习题 1, d)), 或构造对于给定的对称二次型的正交基.

在习题 3 至 8 里, 要求利用习题 2 所指出的方法化二次型为平方和的形式, 取 **Q** 作为基础域, 并且在每一个情形指出导致所要求的结果的坐标变换.

3. $x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3$.

4. $x_1^2 + 5x_2^2 - 4x_3^2 + 2x_1x_2 - 4x_1x_3$.

5. $2x_1^2 + 18x_2^2 + 8x_3^2 - 12x_1x_2 + 8x_1x_3 - 27x_2x_3$.

6. $x_1^2 + 2x_2^2 + x_4^2 + 4x_1x_2 + 4x_1x_3 + 2x_1x_4 + 2x_2x_3 + 2x_2x_4 + 2x_3x_4$.

7. $3x_1^2 + 2x_2^2 - x_3^2 - 2x_4^2 + 2x_1x_2 - 4x_2x_3 + 2x_2x_4$.

8. $3x_1^2 - 2x_2^2 + 2x_3^2 + 4x_1x_2 - 3x_1x_3 - x_2x_3$.

9. 设 f 域 K 上的有限维向量空间 E 上的一个双线性型.

a) 设

$$A = (f(a_i, a_j))_{1 \leqslant i,j \leqslant n}$$

是 f 关于 M 的一个基 $(a_i)_{1 \leqslant i \leqslant n}$ 的矩阵. 把每一个向量 $x \in E$ 与 x 关于所提及的基的坐标组成的列矩阵等同. 证明对于任意 $x, y \in E$ 我们有

$$f(x, y) = {}^t y \cdot A \cdot x.$$

b) 设 B 是 f 关于另一个基 $(b_i)_{1 \leqslant i \leqslant n}$ 的矩阵. 用 P 表示从 (a_i) 到 (b_i) 的过渡矩阵. 证明

$$B = {}^t PAP.$$

c) 由此推出对于所有对称矩阵 $\mathrm{S} \in \mathrm{M}_n(\mathrm{K})$, 存在一个对角矩阵 D 和一个矩阵 $\mathrm{P} \in \mathrm{GL}(n, \mathrm{K})$, 使得

$$\mathrm{S} = {}^t\mathrm{PDP},$$

并且其逆命题成立. 进一步证明如果 K 是代数闭的, 则可以假定 D 的所有对角线元素是 1 或 0, 而如果 $\mathrm{K} = \mathbf{R}$, 可以假定它们等于 $0, 1$ 或 -1.

10. 对于元素在一个代数闭域 K 内的所有 n 阶对称矩阵 S, 存在一个矩阵 $\mathrm{X} \in \mathrm{M}_n(\mathrm{K})$, 使得

$$\mathrm{S} = {}^t\mathrm{XX},$$

其逆命题亦真.

11. 设 f 是交换域 K (配备了一个对合, 参见 §36 的引言) 上的有限维向量空间 E 上的一个 Hermit 型.

a) 设 $\mathrm{A} = (f(a_i, a_j))_{1 \leqslant i,j \leqslant n}$ 是 f 关于 E 的一个基 (a_i) 的矩阵. 证明, 如果把每一个向量 $x \in \mathrm{E}$ 与 x 关于所提及的基的坐标组成的列矩阵等同, 则对于如何 $x, y \in \mathrm{E}$ 有

$$f(x, y) = y^* \cdot \mathrm{A} \cdot x.$$

b) 设 u 是 E 的自同态, 它的关于基 (a_i) 的矩阵是 U, 证明 U 关于 f (假定是非退化的) 的伴随同态关于这个基的矩阵是

$$\mathrm{A}^{-1}\mathrm{U}^*\mathrm{A}.$$

c) 利用对于所有 Hermit 型存在一个正交基的结论, 证明对于所有 Hermit 矩阵 $\mathrm{A} \in \mathrm{M}_n(\mathrm{K})$, 存在一个对角 Hermit 矩阵 $\mathrm{D} \in \mathrm{M}_n(\mathrm{K})$ 和一个可逆矩阵 $\mathrm{P} \in \mathrm{GL}(n, \mathrm{K})$, 使得

$$\mathrm{A} = \mathrm{P}^*\mathrm{DP}.$$

12. 两个 Hermit 矩阵的乘积是一个 Hermit 矩阵, 必须并且只需两个给定的矩阵是交换的.

13. 设 E 是一个有限维复向量空间, 而 f 是 E 上的一个正定 Hermit 型.

a) 证明对于 E 的所有向量子空间 M, 正交投影算子 p_{M} 关于 f 是自伴的.

b) 反之, E 的所有自同态 p, 如果满足关系

$$p = p* = p^2,$$

则存在一个子空间 M, 使得 $p = p_{\mathrm{M}}$.

c) 设 M 和 N 是 E 的两个向量子空间, 又设 M′ (对应的, N′) 是正交于 $\mathrm{M} \cap \mathrm{N}$ 的 $x \in \mathrm{M}$ (对应的, $x \in \mathrm{N}$) 组成的子空间. 证明 p_{M} 和 p_{N} 是交换的, 必须并且只需 M′ 和 N′ 是正交的 (这个状况推广了通常三维空间内两个垂直平面的状况). 如果这个条件满足, 则有

$$p_{\mathrm{M} \cap \mathrm{N}} = p_{\mathrm{M}} \circ p_{\mathrm{N}},$$

$$p_{\mathrm{M}+\mathrm{N}} = p_{\mathrm{M}} + p_{\mathrm{N}} - p_{\mathrm{M}} \circ p_{\mathrm{N}}.$$

d) 能够推广到任意基础域吗? (考虑非迷向子空间.)

14. 设 f 是交换域 K 上的有限维向量空间 E 上的一个对称双线性型. 设 $a_1, \cdots, a_r$ 是 E 的元素, 而 U 是它们生成的子空间. 则下列两个条件是等价的: (i) U 是非迷向的, 并且诸 a_i 组成

U 的一个基; (ii) $f(a_i, a_j)$ 的行列式非零. 推论: 如果 f 不具有任何迷向向量 (例子: $\mathrm{K} = \mathbf{R}$, 并且 f 是正定的), 向量 $a_1, \cdots, a_r$ 是线性无关的, 必须并且只需 $f(a_i, a_j)$ 的行列式非零.

15. 设

$$\mathrm{S} = (a_{ij})_{1 \leqslant i,j \leqslant n}$$

是元素在一个交换域 K 内的对称矩阵. 假定 S 的**主子式**

$$\mathrm{D}_p = \begin{vmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pp} \end{vmatrix} \quad (1 \leqslant p \leqslant n)$$

不是零. 证明存在对角矩阵

$$\mathrm{D} = \begin{pmatrix} c_1 & 0 & 0 & \cdots & 0 \\ 0 & c_2 & 0 & \cdots & \cdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & c_n \end{pmatrix}$$

和形如

$$\mathrm{T} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ t_{21} & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ t_{n1} & t_{n2} & t_{n3} & \cdots & t_{n,n-1} & 1 \end{pmatrix}$$

的三角矩阵, 使得

$$\mathrm{S} = \mathrm{T}\mathrm{D}\,{}^t\mathrm{T}.$$

证明其中的 D 和 T 由 S 完全确定, 并且 D 的对角线元素由下列关系给定:

$$c_1 = \mathrm{D}_1, c_2 = \mathrm{D}_2/\mathrm{D}_1, \cdots, c_n = \mathrm{D}_n/\mathrm{D}_{n-1}.$$

推论: 设

$$q(x) = \sum a_{ij} x_i x_j \quad (a_{ij} = a_{ji})$$

是一个二次型, 矩阵 (a_{ij}) 的主子式不是零, 则借助由一个三角矩阵给定的坐标变换可以化成平方和.

(为了证明 T 和 D 的存在性, 可以利用 §23 习题 11, 或进行关于 n 的归纳推理, 写出

$$\mathrm{S} = \begin{pmatrix} \mathrm{S}_1 & {}^t u \\ u & a_{nn} \end{pmatrix},$$

其中 S_1 是一个 $n-1$ 阶方阵, 而 u 是一个 $n-1$ 个元素的行矩阵, 并且对于 D 和 T 利用一个类似的分块分解.)

¶16. 保留上题的记号, 假定 S 的秩 r 是任意的, 我们要把 S 表示成

$$\mathrm{S} = \mathrm{T}\mathrm{D}\,{}^t\mathrm{T},$$

其中的 $c_1 \neq 0, \cdots, c_r \neq 0, c_{r+1} = \cdots = c_n = 0$. 证明问题具有一个解, 必须并且只需

$$\mathrm{D}_p \neq 0 \quad 对于\ 1 \leqslant p \leqslant r.$$

17. 设 E 是交换域 K 上的有限维向量空间, 而 f 是 E 上的非退化的对称双线性型; 用 q 记二次型 $f(x,x)$. 假定存在对于 f 的非零迷向向量. 证明这时对于所有 $c \in \mathrm{K}$ 存在 $x \in \mathrm{E}$, 使得

$$q(x) = c.$$

¶18. 设 E 是交换域 K 上的 n 维向量空间, 而 f 是 E 上的非退化的对称双线性型.

a) 设 H 是对于 f 非迷向的一个向量子空间. 证明存在 f 的唯一的自同构 s_{H}, 使得

$$s_{\mathrm{H}}(x) = \begin{cases} x, & 如果\ x \in \mathrm{H}, \\ -x, & 如果\ x \in \mathrm{H}^{\perp} \end{cases}$$

(s_{H} 称为**关于 H 的对称**).

b) 设 u 是 f 的一个自同构, 而 x 是对于 f 的一个非迷向向量. 证明向量 $u(x)+x, u(x)-x$ 至少有一个是非迷向的. 由此推出在 E 内存在一个非迷向子空间 H, 使得 $s_{\mathrm{H}}(u(x)) = x$ [取 H 为正交于 $u(x)-x$ 的子空间, 或由 $u(x)-x$ 生成的子空间].

c) 用关于 n 的归纳推理由此推出 f 的所有自同构是至多 n 个关于 E 的非迷向子空间的对称的乘积.

¶¶19. 设 E 是交换域 K 上的有限维向量空间, 而 f 是 E 上的对称双线性型. 称 E 的一个子空间 U 对于 f 是**完全迷向的**, 如果对于所有 $x \in \mathrm{U}$ 有 $f(x,x)=0$. 称 U 是**极大**完全迷向的, 如果它不包含于另外的对于 f 完全迷向子空间内. 我们打算证明 f 的所有极大完全迷向的子空间有同样的维数, 称为 f 的 (或对应的二次型的) **指标**.

a) U 对于 f 是完全迷向的, 必须并且只需

$$f(x,y) = 0 \quad 对于任意\ x, y \in \mathrm{U},$$

即 $\mathrm{U} \subset \mathrm{U}^{\perp}$ (利用习题 1).

b) 设 U 和 V 是两个对于 f 的完全迷向的子空间. 证明对于所有 $x \in \mathrm{U} \cap \mathrm{V}^{\perp}$, 子空间 $\mathrm{V} + \mathrm{K}x$ 是完全迷向的.

c) 设 U 和 V 是两个对于 f 的完全迷向的子空间, 再设 M 是 $\mathrm{U} \cap \mathrm{V}$ 在 U 内的补空间, 而 N 是 $\mathrm{U} \cap \mathrm{V}$ 在 V 内的补空间. 证明

$$\mathrm{U} \cap \mathrm{V}^{\perp} = (\mathrm{U} \cap \mathrm{V}) \oplus (\mathrm{M} \cap \mathrm{N}^{\perp}).$$

证明 $\mathrm{M} \cap \mathrm{N}^{\perp}$ 的元素通过解 $r = \dim(\mathrm{M})$ 个未知元 $s = \dim(\mathrm{N})$ 个齐次线性方程组的解而得到. 由此得到, 如果

$$\dim(\mathrm{V}) < \dim(\mathrm{U}),$$

则存在一个不在 V 内的 $x \in \mathrm{U}$, 使得 $\mathrm{V} + \mathrm{K}x$ 是完全迷向的.

d) 证明所有完全迷向子空间包含于至少一个极大迷向子空间内. 借助问题 c) 由此推出题目开头所宣布的结果.

¶¶20. (Witt 定理的证明) 设 K 是特征异于 2 的交换域, E 是 K 上的有限维向量空间, 而 f 是 E 上的非退化对称双线性型. 设 M 和 N 是 E 的两个同样维数的子空间, 而 u 是从 M 到 N 上的双射的线性映射. 我们打算证明以下两个性质等价: (i) 存在 f 的一个同构在 M 上与 u 重合; (ii) 对于任意 $x, y \in \mathrm{M}$ 有 $f[u(x), u(y)] = f(x, y)$. 由于 (i) 蕴含 (ii) 是平凡的, 下面我们限于证明 (ii) 蕴含 (i).

a) 设 x 和 y 是 E 的两个元素, 使得

$$f(x,x) = f(y,y) \neq 0,$$

证明存在 f 的一个把 x 映射到 y 的同构 (证明 $x-y$ 和 $x+y$ 不都是迷向的, 取关于 H 的对称, 这里 H 是垂直于 $x-y$ 的平面, 或是垂直于 $x+y$ 的平面).

b) 假定 M 和 N 不是完全迷向的. 借助问题 a) 证明可以假定对于一个非迷向的 $x \in \mathrm{M}$ 有 $u(x) = x$. 设 E′ 是垂直于 x 的平面. 证明为了构造延拓 u 的 f 的一个自同构, 只需构造 f 在 E′ 的限制 f' 的在 $\mathrm{E}' \cap \mathrm{M}$ 上等于 u 的自同构. 由此通过关于 dim(E) 的归纳推理推出在这种情形的 Witt 定理.

c) 以下假定 M 和 N 是完全迷向的. 选择一个 $x \notin \mathrm{M}^{\perp}$. 证明一个 $y \notin \mathrm{N}^{\perp}$, 使得

$$f(y, u(z)) = f(x, z) \quad \text{对于所有 } z \in \mathrm{M}.$$

证明还可以假定

$$f(y,y) = f(x,x)$$

(适当选择 $t \in \mathrm{K}, n \in \mathrm{N}$, 用 $y + tn$ 代替 y).

d) 由 c) 推出存在非完全迷向的子空间 $\mathrm{M}' \supset \mathrm{M}$ 和 $\mathrm{N}' \supset \mathrm{N}$, 以及从 M′ 到 N′ 上的一个自同构 u', 使得

$$f(u'(x), u'(y)) = f(x,y) \quad \text{对于任意的 } x, y \in \mathrm{M}',$$

$$u' = u \text{ 在 M 上},$$

由此出发完成 Witt 定理的证明.

¶¶21. 设 E 是交换域 K 上的一个有限维向量空间, f 是 E 上的一个非退化对称双线性型, 而 M 是 E 的维数为 r 的一个完全迷向的子空间.

a) 证明在 E 内存在不正交于 M 的非迷向的向量.

b) 证明存在 E 的一个完全迷向的子空间 N, 使得

$$\mathrm{E} = \mathrm{M}^{\perp} \oplus \mathrm{N}$$

(取 M 的关于由上一个问题中构造的一个向量生成的直线的对称).

c) 设 $\mathrm{H} = \mathrm{M}^{\perp} \cap \mathrm{N}^{\perp}$, 证明

$$\mathrm{E} = \mathrm{M} \oplus \mathrm{H} \oplus \mathrm{N},$$

并且如果 M 是极大完全迷向的, 则 H 不含有任何非零迷向向量. 合并 M,H 和 N 的基以组成 E 的一个基, 证明 f 关于这个基的矩阵有形式

$$\mathrm{S} = \begin{pmatrix} 0 & 0 & \mathrm{A} \\ 0 & \mathrm{S}_1 & 0 \\ {}^t\mathrm{A} & 0 & 0 \end{pmatrix},$$

其中的 A 是 r 阶可逆的方阵, 而 S_1 是 $n-2r$ 阶的对称矩阵.

d) 找出一个形如

$$\begin{pmatrix} U & 0 & 0 \\ 0 & V & 0 \\ 0 & 0 & W \end{pmatrix}$$

的矩阵 (U 和 W 是 r 阶方阵, 而 V 是 $n-2r$ 阶的方阵) 表示 f 的自同构关于 E 在 c) 内所考虑的基的矩阵. 由此推出对于向量空间 M 的所有自同构, 存在 f 的一个在 M 上化为 u 的自同构. 能够从 Witt 定理得到这个结果吗?

22. 设 $q(x)$ 是有限维实 (对应的, 复) 向量空间 E 上的一个二次型, 证明存在 E 的一个基, 使得关于这个基 q 有形式

$$x_1^2+\cdots+x_p^2-(x_{p+1}^2+\cdots+x_{p+q}^2) \tag{$*$}$$

(对应的,

$$x_1^2+\cdots+x_r^2).$$

对于习题 3 至 8 中的二次型 (在实的情形和复的情形) 求这样的基.

23. 设 $f_1,\cdots,f_{p+q}$ 是有限维实空间 U 上的线性型. 假定在 U 上二次型

$$q(x)=f_1(x)^2+\cdots+f_p(x)^2-f_{p+1}(x)^2-\cdots-f_{p+q}(x)^2$$

是正定的, 即满足

$$q(x)>0 \quad \text{对于所有 } x\neq 0.$$

证明有

$$\dim(\mathrm{U})\leqslant p.$$

(注意在相反的情形, 存在 $x\neq 0$, 使得 $f_1,\cdots,f_p$ 全是零.)

¶24. 设 $q(x)$ 是有限维实向量空间 E 上的一个二次型. 选择 E 的一个基, 使得关于这个基 q 有习题 22 $(*)$ 的形式. 借助上一个习题证明 E 的 $q(x)$ 在其上是正定的所有子空间 U 的维数至多是 p. 由此推出 (二次型的**惯性定律**) 习题 22 中的 p 和 q 不依赖 E 的基的选择, 关于这个基 $q(x)$ 化成形式 $(*)$.

[由 $(*)$ 式中正平方的数目 p 和负平方的数目 q 组成的序偶 (p,q) 称为所考虑的二次型或对应的对称双线性型的**符号**. 数 p 是在其上给定的二次型是正定的子空间的最大维数, 而 q 则是在其上给定的二次型是负定的子空间的最大维数.]

¶25. 交换域 K 上的有限维向量空间 E 上的两个二次型 q 和 q' 是**等价的**, 如果存在 E 的一个自同构 u, 使得

$$q'(x)=q(u(x)) \quad \text{对于所有的 } x\in \mathrm{E}.$$

假定基础域是 $\mathbf{R}$, 证明 q 和 q' 是等价的, 必须并且只需它们有同样的符号.

26. 证明 $\mathbf{R}^3$ 上的两个二次型是等价的, 并且构造把第一个变换到第二个的 $\mathbf{R}^3$ 的一个自同构:

$$2x^2+9y^2+3z^2+8xy-4xz-10yz,$$

$$3x^2+3y^2+6z^2-48xy-4xz+8yz.$$

对于二次型

$$5x^2 + 5y^2 + 2z^2 + 8xy + 6xz + 6yz \quad 和 \quad 4x^2 + y^2 + 9z^2 - 12xz$$

解答同样的问题

¶27. 考虑 $\mathbf{R}^n$ 上的对应于二次型

$$q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$$

的对称双线性型 f, 并且假定 $p \leqslant q$. 证明由向量

$$e_1 + e_{p+1}, e_2 + e_{p+2}, \cdots, e_p + e_{2p}, e_{p+q+1}, \cdots, e_n$$

(其中 $e_1, \cdots, e_n$ 是 $\mathbf{R}^n$ 的典范基) 生成的子空间对于 f 是极大完全迷向的 (习题 19).

由此推出如果 f 是维数为 n 的实向量空间上的符号为 (p, q) 的对称双线性型, 则 f 的指标 (习题 14) 是 $r + n - p - q$, 其中的 r 是两个整数 p 和 q 当中较小的一个.

Lorentz 型极大完全迷向子空间的维数是多少?

¶28. 设 f 是有限维实向量空间 E 上的一个非退化对称双线性型. 设 M 和 N 是 E 的两个向量子空间, $\dim(\mathrm{M}) = \dim(\mathrm{N})$. 存在映射 M 到 N 上的 f 的一个自同构, 必须并且只需 f 到 M 和 N 的限制有同样的符号 (利用 Witt 定理和习题 24, 25).

取 f 是 Lorentz 型, 并且考虑在 E 的所有向量子空间的集合上的等价关系: "存在 f 的一个自同构 u, 使得 $u(\mathrm{M}) = \mathrm{N}$". 商集合有多少元素?

对于 $\mathbf{R}^5$ 上的二次型

$$x^2 + y^2 + z^2 - t^2 - u^2$$

解答同样的问题.

29. 一个复 Hermit 矩阵 $\mathrm{H} \in \mathrm{M}_n(\mathbf{C})$ 是正定的, 必须并且只需存在一个矩阵 $\mathrm{X} \in \mathrm{M}_n(\mathbf{C})$, 使得

$$\mathrm{H} = \mathrm{X}^*\mathrm{X}.$$

如果 H 进而是实的, 则可以假定 X 是实的 [利用习题 11, c), 并且注意到如果 H 是半正定的, 则 H 的对角线元素是非负的].

由此推出如果

$$\mathrm{H} = (a_{ij})_{1 \leqslant i,j \leqslant n}$$

是半正定复 Hermit 矩阵, 则有

$$\mathrm{D}_p = \begin{vmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pp} \end{vmatrix} \geqslant 0 \quad 对于\ 1 \leqslant p \leqslant n,$$

并且

$$\mathrm{D}_p = 0 \quad 蕴含 \quad \mathrm{D}_{p+1} = \cdots = \mathrm{D}_n = 0$$

(首先证明 $\mathrm{D}_n \geqslant 0$, 然后用行列式为 D_p 的矩阵代替 H, 并且利用习题 16). 情形 $n = 2$ 如何? (重新得到 Cauchy-Schwarz 不等式和 "三项式符号".)

¶30. 设

$$H = (h_{ij})_{1 \leqslant i,j \leqslant n}$$

是一个半正定的复 Hermit 矩阵. 证明存在一个复三角矩阵

$$T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & t_{nn} \end{pmatrix},\ \text{其中 } t_{ii} \text{ 是非负的,}$$

使得

$$H = T^* T$$

(利用习题 15 和 16). 如果 H 是可逆的, 则 T 是唯一的. 由前述内容推出对于所有半正定 Hermit 矩阵 H 有不等式

$$0 \leqslant \det(H) \leqslant h_{11} h_{22} \cdots h_{nn}.$$

如果 H 是可逆的, 进而不能有等式, 除非 H 是对角矩阵.

¶31. 设 A 是复元素的可逆的方阵. 证明存在一个酉矩阵 U 和一个三角矩阵 $T = (t_{ij})$, 对于所有的 i 有 $t_{ii} > 0$, 使得

$$A = U \cdot T,$$

并且 U 和 T 是由 A 唯一确定的 (利用上一个习题到 A^*A). 证明如果 A 是实矩阵, 则 U 和 T 是实矩阵.

¶32. 设 $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ 是一个复矩阵, 证明

$$|\det(A)|^2 \leqslant \prod_{j=1}^{n} (|a_{1j}|^2 + \cdots + |a_{nj}|^2).$$

¶33. 一个复 Hermit 矩阵是正定的, 必须并且只需它的所有主子式是正的 (利用习题 30, 并且通过给定矩阵的主子式计算 t_{ii}).

34. 借助上一个习题, 求使得下列二次型是正定的 t 的实值:

$$\begin{aligned} &5x_1^2 + x_2^2 + tx_3^2 + 4x_1x_2 - 2x_1x_3 - 2x_2x_3, \\ &2x_1^2 + x_2^2 + 3x_3^2 + 2tx_1x_2 + 2x_1x_3, \\ &2x_1^2 + 2x_2^2 + x_3^2 + 2tx_1x_2 + 6x_1x_3 + 2x_2x_3, \\ &t(x_1^2 + x_2^2 + x_3^2) + 4x_1x_2 + 6x_1x_3 + 8x_2x_3. \end{aligned}$$

¶¶35. 证明所有半正定 Hermit 矩阵是形如

$$\begin{pmatrix} c_1\bar{c}_1 & \cdots & c_1\bar{c}_n \\ \vdots & & \vdots \\ c_n\bar{c}_1 & \cdots & c_n\bar{c}_n \end{pmatrix}$$

的矩阵的和, 并且其逆命题亦真.

由此推出如果两个复矩阵 $(a_{ij})_{1\leqslant i,j\leqslant n}$ 和 $(b_{ij})_{1\leqslant i,j\leqslant n}$ 是半正定 Hermit 的, 则给定的两个矩阵的相同指标的元素相乘得到的矩阵

$$(a_{ij}b_{ij})_{1\leqslant i,j\leqslant n}$$

也是半正定 Hermit 的.

36. 设 H 是一个复 Hermit 矩阵. 证明 $1-i\mathrm{H}$ 是可逆的,

$$\mathrm{U}=(1+i\mathrm{H})(1-i\mathrm{H})^{-1}=(i-\mathrm{H})(i-\mathrm{H})^{-1}$$

是酉矩阵, 并且 -1 不是 U 的特征值. 反之, 所有的 -1 不是其特征值的酉矩阵可以以这种方式得到 (Cayley 变换).

¶37. 设 H 是一个半正定复 Hermit 矩阵. (通过借助一个酉矩阵把 H 化成对角矩阵) 证明存在唯一的一个半正定 Hermit 矩阵 H′, 使得

$$\mathrm{H}=\mathrm{H}'^2$$

(称 H′ 是 H 的半正定平方根, 并且写作 $\mathrm{H}'=\mathrm{H}^{\frac{1}{2}}$).

¶38. 设 S 和 T 是两个复 Hermit 矩阵, 假定 S 是正定的. 证明 ST 的特征值是实数, 如果 T 是半正定的, 则 ST 的特征值是非负实数 (利用习题 37).

¶39. 设 A 是复元素的可逆的方阵, 证明矩阵

$$\mathrm{A}^*\mathrm{A}$$

是正定 Hermit 矩阵. 通过考虑它的非负平方根, 证明可以写出

$$\mathrm{A}=\mathrm{U}\cdot\mathrm{H},$$

其中 U 是酉矩阵, 而 H 是正定 Hermit 矩阵; 进而证明问题仅有一个解. 证明如果 A 是实的, 则 U 和 H 也是实的.

¶¶40. 设 $\mathrm{F}\subset\mathrm{M}_n(\mathbf{C})$ 是两两可以交换的矩阵的一个集合. 证明存在一个酉矩阵 $\mathrm{U}\in\mathrm{M}_n(\mathbf{C})$, 使得对于所有 $\mathrm{A}\subset\mathrm{F}$ 矩阵 UAU^{-1} 是对角矩阵. 几何解释如何? (利用 §34 的习题 22.)

在习题 41 至 47 里要求借助一个实正交矩阵把给定的实对称矩阵化成一个对角矩阵.

41. $\begin{pmatrix}6&-2&2\\-2&5&0\\2&0&7\end{pmatrix}$. 42. $\begin{pmatrix}11&8&2\\8&5&-10\\2&-10&2\end{pmatrix}$. 43. $\begin{pmatrix}8&4&-1\\4&-7&4\\-1&4&8\end{pmatrix}$.

44. $\begin{pmatrix}17&-2&-2\\-2&14&-4\\-2&-4&14\end{pmatrix}$. 45. $\begin{pmatrix}0&1&-3&0\\1&0&0&-3\\-3&0&0&1\\0&-3&1&0\end{pmatrix}$.

46. $\begin{pmatrix}5&-5&1&3\\-5&5&3&1\\1&3&5&-5\\3&1&-5&5\end{pmatrix}$. 47. $\begin{pmatrix}9&0&0&0\\0&5&4&-2\\0&4&5&2\\0&-2&2&8\end{pmatrix}$.

¶48. 设 E 是有限维实向量空间, f 是 E 上的正定对称双线性型, 而 u 是 f 的一个自同构. 假定 E 仅有的在 u 下稳定的向量子空间是 $\{0\}$ 和 E. 证明则有下列两个状况之一: (i) E 是一维的, 而 u 等于 1 或 -1; (ii) E 是二维的, 并且存在一个对于 f 正交的基, 而 u 关于这个基的矩阵有形式

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

其中的 φ 不是 π 的整倍数的一个实数.

¶49. 设 E 是有限维实向量空间, f 是 E 上的正定对称双线性型, 而 u 是 f 的一个自同构. 证明存在 E 的一个相关于 f 的正交的基, 并且 u 关于这个基的矩阵的形式是

$$\begin{pmatrix}
1 & & & & & & & & & & & \\
& \ddots & & & & & & & & & & \\
& & 1 & & & & & & & 0 & & \\
& & & -1 & & & & & & & & \\
& & & & \ddots & & & & & & & \\
& & & & & -1 & & & & & & \\
& & & & & & \cos\varphi_1 & -\sin\varphi_1 & & & & \\
& & & & & & \sin\varphi_1 & \cos\varphi_1 & & & & \\
& & & & & & & & \ddots & & & \\
& & 0 & & & & & & & & & \\
& & & & & & & & & & \cos\varphi_h & -\sin\varphi_h \\
& & & & & & & & & & \sin\varphi_h & \cos\varphi_h
\end{pmatrix},$$

其中的 φ_i 为不是 π 的整倍数的一个实数.

50. 对于所有实数 t, 令

$$\mathrm{U}(t) = \begin{pmatrix} \cos t & \sin t & 0 \\ -\sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathrm{V}(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{pmatrix}.$$

证明, 对于所有行列式为 1 的三阶实正交矩阵 X (通常空间内的旋转矩阵) 存在实数 φ, ψ, θ, 使得

$$\mathrm{X} = \mathrm{U}(\varphi)\mathrm{V}(\theta)\mathrm{U}(\psi).$$

(注意到如果有两个规范正交基 i, j, k 和 u, v, w, 从第一个过渡到第二个可以这样实现, 首先绕 k 旋转把 i 变到 u 和 v 生成的平面, 再绕平面 ij 和 uv 的交线旋转把 k 变到 w, 最后绕 w 旋转.) 数 φ, ψ 和 θ 称为矩阵 X (或对应旋转的) 的 **Euler 角**.

¶¶51. 考虑符号为 (p, q) 的 $p+q$ 阶 Hermit 矩阵

$$\mathrm{S} = \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix}.$$

证明 S 的自同构群, 即满足关系

$$W^*SW = S$$

的矩阵 $W \in GL(p+q, \mathbf{C})$ 的群, 是矩阵

$$W = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

的集合, 它们具有下列性质: 存在 p 列 q 行的矩阵 Z, 使得

$$1_p - Z^*Z$$

是正定的, 还存在 p 阶和 q 阶的酉矩阵 U 和 V, 使得有关系

$$A = (1 - Z^*Z)^{-\frac{1}{2}}U, \quad B = Z^*(1 - Z^*Z)^{-\frac{1}{2}}V,$$
$$C = Z(1 - Z^*Z)^{-\frac{1}{2}}U, \quad B = (1 - Z^*Z)^{-\frac{1}{2}}V,$$

并且 U,V 和 Z 由 W 完全确定. 如果 W 是实的, 则 U,V 和 Z 也是实的. [注意: 当 H 是正定时我们令

$$H^{-\frac{1}{2}} = (H^{\frac{1}{2}})^{-1}.]$$

¶52. 设 V 是交换域 K 上的偶数 $n = 2m$ 维向量空间.

a) 设 f 是 V 上的非退化交错 (§22) 双线性型. 取两个向量 a 和 b, 使得 $f(a,b) = 1$, 并且考虑由 a 和 b 生成的平面 $P \subset V$. 证明 f 在 P 上的限制是非退化的, 由此推出 V 是 P 和对于 f 垂直于 P 的向量的子空间 V' 的直和. 由此得到 V 具有一个这样的基, 对于这个基 $f(x,y)$ 的通过 x 和 y 的坐标的表达式是

$$f(x,y) = \sum_{i=1}^{m} (x_i y_{2n-i} - y_i x_{2n-i}).$$

由此推出 V 上的两个非退化的交错的双线性型可以通过 V 的自同构从一个变换到另一个. 在奇数维空间上是否存在非退化的交错的双线性型?

b) 选择 V 上的一个非退化的交错的双线性型, V 的所有二维子空间 P, 如果 f 在 P 上不恒等于零, 则称为 V 的**非退化平面**. 设 Sp(V) 是 V 的这样的自同构 u 的群 ("对称群"), 这些 u 使得

$$f(u(x), u(y)) = f(x,y) \quad \text{对于任意 } x, y \in V.$$

证明 Sp(V) 传递地作用在非退化平面的集合 X 上, 并且如果 $P \in X$, 让 P 不动的 Sp(V) 的子群 (即使得 $u(P) = P$ 的 $u \in Sp(V)$ 的集合) 同构于乘积

$$SL(2, K) \times \mathfrak{S}_p(W),$$

其中 W 是对于 f 的 P 在 V 内的正交子空间.

c) 现在假定 K 是 q 个元素的有限交换域. 设 x 是 V 的一个非零元. 证明 x 属于 q^{n-2} 个非退化平面. 由此推断出公式

$$\mathrm{Card}(X) = \frac{q^n - 1}{q^2 - 1} q^{n-2}.$$

设 P 是一个非退化平面, 而 W 是在 V 内的 P 的正交子空间. 证明

$$\mathrm{Card}(\mathrm{Sp}(V)) = (q^n - 1)q^{n-1}(\mathrm{Sp}(W)).$$

由此通过关于 n 的归纳推理推断出公式

$$\mathrm{Card}(\mathrm{Sp}(V)) = q^{m(2m+1)} \prod_{i=1}^{n} (1 - 1/q^{2i}).$$

¶53. 设 X 是有 6 个元素的一个集合. 设 Y 是 X 的有 0 个元素或 2 个元素的子集的集合. 在 Y 上定义一个对称的运算如下:

$$\begin{aligned}
&A + \varnothing = A \quad \text{对于所有的 } A \in Y,\\
&A + A = \varnothing \quad \text{对于所有的 } A \in Y,\\
&A + B = A \cup B - A \cap B, \quad \text{如果 } A \cap B \text{ 有一个元素},\\
&A + B = X - (A \cup B), \quad A \text{ 和 } B \text{ 是不交的并且是非空的}.
\end{aligned}$$

a) 证明这个运算使得 Y 成为一个 16 阶的交换群, 并且中性元为 $\varnothing$. 证明 Y 可以 (用唯一的方式) 配备在域 $K = \mathbf{Z}/2\mathbf{Z}$ 上的一个四维的向量空间结构.

b) 如果 $A, B \in Y$, 用 $f(A, B)$ 表示由同余关系

$$f(A, B) \equiv \mathrm{Card}(A \cap B) (\mathrm{mod}\ 2)$$

定义的 K 的元素. 证明 f 是 Y 上的一个非退化的交错的双线性型.

c) 设 A, B, C 是具有两个元素的互不相交的 X 的三个子集. 证明 $\{\varnothing, A, B, C\}$ 是 (关于 f 的) Y 的一个二维的完全迷向的子空间.

d) 设 S_X 是 X 的置换群, 这个群同构于对称群 S_6. 再设 Sp(Y) 是保持型 f 的 Y 的自同构群 (这个群经常记作 $\mathrm{Sp}_4(F_2)$). 证明 S_X 的所有元素定义 Sp(Y) 的一个元素, 并且这样得到的同态 $\varepsilon: S_X \to \mathrm{Sp}(Y)$ 是一个同构. (首先证明 ε 是单射的, 然后比较 S_X 和 Sp(Y) 的阶.)

e) 设 F 是使得 $\sum_{x \in X} f(x) = 0$ 的函数 $f: X \to K$ 的向量空间. 设 $V = F/\{0, 1\}$ 是 F 关于由常函数 1 生成的一维子空间的商空间. 令所有 $A \in Y$ 对应特征函数 θ_A (在 A 上等于 1, 在 X − A 上等于 0). 证明 $A \to \theta_A$ 定义从向量空间 Y 到向量空间 V 的一个同构. 证明这个同构把 b) 的双线性型变换为

$$u(\theta, \theta') = \sum_{x \in X} \theta(x)\theta'(x).$$

参考文献

[1] A. A. Albert, *Introduction to Algebraic Theories*, University of Chicago Press, 1941.

多项式, 矩阵, 行列式, 线性方程组, 矩阵的化简 (包括 Jordan 标准形), 在环和域的推广. 经典的初等教科书, 十分集中和严谨的陈述. 习题.

[2] A. A. Albert, *Foundamental Concepts of Higher Algebra*, University of Chicago Press, 1956.

群, 环, 域, 向量空间和矩阵, 交换域的代数扩张, 有限域. 比前一个著作水平明显更高, 并且面向更高等的学生. 习题.

[3] E. Artin, *Geometric Algebra*, Interscience Publishers, New York, 1957.

[4] E. Artin, *Algèbre Géométrique*, Gauthier-Villars, Paris, 1962.

向量空间, 对偶, 线性方程组, 群, 域 (特别快的陈述); 初等几何的 (仿射和射影) 公理结构; 正交对称, 和线性群的几何研究. 主要用以补充本书的 §36. 从极其初等的水平开始的直达尚未解决的问题的一流著作.

[5] G. Birkhoff and S. Maclane, *A Survey of Modern Algebra*, Macmillan, New York, 1941.

涵盖了本书几乎所有主题, 此外陈述了交换域的代数扩张理论 (Galois 理论, 等等). 容易接受的和现代化的陈述, 带有大量的一般说来十分简单的习题.

[6] N. Bourbaki, *Eléments d'histoire des mathématiques*, Hermann, Paris, 1960.

对于那些仅对数学的实际方面感兴趣的人来说, 阅读这个由 N. Bourbaki 在他们的各卷《数学论著》中附加的 “历史评注” 的汇编显然不是必须的. 但是对于那些渴望了解为什么要引进这里所陈述的无论是古典的还是现代的概念的人来说, 最好的做法是研究这些历史的评注. 阅读相当困难, 作者不限于用基础法语书写它们的评注.

[7] I. M. Gel'fand, *Leksii po lineǐnoǐ algebre*, Gostekhizdat, Moscow, 1951.

[8] I. M. Gel'fand, *Lectures on Linear Algebre*, Interscience Publishers, New York, 1961.

本书主要是以更详尽的方式陈述我们书的 §34, §35 和 §36 中所处理的主题 (尤其是关于 Jordan 理论), 假定读者已经熟知行列式和线性方程组理论. 本书显然有意要作为 Hilbert 空间理论和典型群 (正交的, 酉的, 线性的, 等等) 理论的引论. 人们可能会指责这本书, 一方面没有充分区分纯线性的性质和涉及使用二次型的性质, 另一方面忽略了 **R** 和 **C** 之外的域. 极其清晰的风格.

[9] W. Greub, *Linear Algebra*, Springer, Berlin, 1958.

涵盖我们书的 §10 至 §25 和 §34 至 §36, 某些环节 (尤其是交错多重线性型) 更详尽. 极好的陈述, 世界最精美的印刷. 英文第二版.

[10] P. R. Halmos, *Finite-dimentional Vector Spaces*, Van Nostrand, Princeton, 1958.

向量空间, 对偶, 商空间, 张量积和交错多重线性型, 线性映射和矩阵, 特征值, Jordan 标准形, Euclid 空间和 Hermit 矩阵. 这本书 (首次出版于 1942 年) 明显地旨在作为 Hilbert 空间理论的引论 (半正定二次型的研究几乎占了三分之一的篇幅), 而宁可忽视了其他重要的方面 (线性方程组, 行列式, 初等因子). 初学者比较易于接受的十分简单和非常几何的陈述. 理论的习题一般十分容易.

[11] H. Hasse, *Höhere Algebra*, 2 vols., Walter de Gruyter (Sammlung Göschen), Berlin, 1951–1957.

[12] H. Hasse, *Higher Algebra*, 2 vols., F. Ungar, New York, 1954.

第一卷讲述环, 域, 多项式, 群, 线性方程组, (域上的) 和行列式理论; 第二卷讲述多项式除法, 交换域的代数扩张和 Galois 理论. 由于作者是 "抽象" 代数的奠基者之一, 这本著作 (第一版在 1926—1927 年问世) 不仅至今尚未过时, 而且甚至许多观点上仍然比自那个年代以来乃至近年来出版的许多著作更近代; 堪称典范的明晰和严格的风格. 仅有的不足在于由于受 Göschen 丛书篇幅的限制内容过分浓缩. 美国的翻译本也未能克服这一不足, 尽管至少可以使用疏密正常的印刷.

[13] K. Hoffman and R. Kunze, *Linear Algebra*, Prentice-Hall, Englewood Cliffs, 1961.

近年来出现在美国的众多初等著作中的一本. 这些书的大部分的数学趣味多少有些令人怀疑, 不过这一本由于其叙述的严格和风格的端正远胜于中等水平. 这本书初始是面向 M. I. T 的本科生, 并且法国的初学者完全能够懂得. 为数众多的习题. 涵盖我们的书的 §10 至 §24 和 §34 至 §36.

[14] A. G. Kuroš, *Kurs vysšeǐ algebra*, Fizmatgiz, Moscou, 1959.

涵盖我们的书的所有章节 (除去集合论), 并且在某些主题 (d'Alembert 定理的证明, 实系数代数方程的实根, 根的对称函数和消元法, 有限交换群的结构, 等等) 相当深入. 叙述整体上是经典的, 而近代的概念 (向量空间, 域, 群) 是在实际上不需要时引入的; 一些人会

发现是相当惬意的.

[15] S. Lang, *Algebra*, Addison-Wesley, Reading, 1965.

群, 环, 域, 多项式, Noether 模; 域论 (代数扩张和超越扩张, Galois 理论, 赋值环); 双线性型, Jordan 标准形, 张量代数, 有限群表示. 虽然 Lang 从十分初等的水平出发, 但其陈述显然针对这样的读者: 他们掌握了我们书的基本内容, 希望更深入, 那么 Lang 的书使得他们能够以极快的速度前进. 许多有益的习题.

[16] A. Lentin and J. Rivaud, *Leçons d'Algèbre moderne*, Vuibert, Paris, 1961.

涵盖了这里所讲的内容的本质部分; 特别是面向数学专门化的学生, 但是对于一般数学课程的学生同样适用, 因为代数的这两种教学课程十分靠近. 许多的习题.

[17] A. Lichnérowicz, *Algèbre et Analyse linéaires*, Masson, Paris, 19–.

向量空间, 矩阵计算, 线性方程组, 行列式和交错型, 特征值, 以及并非针对初学者的分析的一些章节. 缺少例子, 还缺少习题.

[18] A. I. Mal'cev, *Foundation of Linear Algebra*, W. H. Freeman, San Francisco, 1963.

矩阵和行列式, 向量空间, Jordan 标准形, Euclid 空间, 二次型, 正交、对称、酉变换, 等等, 多重线性型, 张量, 外代数. 习题. 极好的陈述, 尽管初学者也许不能立刻领会, 但是他会多年受用.

[19] L. Mirsky, *An Introduction to Linear Algebra*, Clarendon Press, Oxford, 1955.

涵盖我们的书的 §10 至 §24 和 §34 至 §36, 某些主题更加详尽 (尤其是 Jordan 标准形), 叙述完备而且恰当, 还有些古怪的细节 (例如引进 "向量的空间" 是 K^n 的子集, 还有 "线性流形" 是本书以及所有数学家的向量空间). 不过是非常好的参考书. 十分多的并且有趣的习题. 浓重的不列颠风格和印刷.

[20] G. D. Mostow, J. H. Sampson and J. P. Meyer, *Fundamental Structures of Algebra*, McGraw Hill, New York, 1963.

内容和水平与我们的书相当接近, 关于线性常微分方程和张量有些补充, 可能对于初学者更容易入门. 叙述非常清晰, 带有许多例子和习题. 当然除了我们的书, 这本整体上引论性的代数书, 在目前世界图书市场上能够买到的书中我们认为最好的.

[21] O. Schreier and E. Sperner, *Einführung in die analytische und Algebra*, 2 vols. Vandenhoeck und Rupprecht, Göttingen, 1955.

[22] O. Schreier and E. Sperner, *Introduction to Modern Algebra and Matrix Theory*, Chelsea Publishing Coy., New York, 1951.

向量空间 $\mathbf{R}^n$, 线性方程, 行列式和交错多重线性型, 二次型和位移, 域, 多项式, d'Alembert 和 Gauss 定理, 有限生成交换群的构造, 矩阵, 特征值, Jordan 标准形. 此外, 德国版有关于射影几何的长长的一章, 而美国译本则没有. O. Schreier 和 E. Sperner 的书的第一版在 1931 年出版, 在那个时代, 还有前面提到的 Hasse 的小书, 是最早系统使

用 “几何” 方法进行初等陈述, 并且深受 19 世纪 20 年代德国数学学派的影响. 关于 Hassel 的书, 我们说它是一本比随后出版的绝大多数著作更近代的著作, 同样适用于 O. Schreier 和 E. Sperner 的书. 比 Hasse 的书更可读的风格 (不刻意节省纸张). 推荐阅读. 大量的习题.

[23] G. E. Šilov, *Vvedenie v teoriyu lineĭnykh prostranstv*, Gostekhizdat, Moscow, 1956.

[24] G. E. Shilov, *An Introduction to the Theory of Linear Spaces*, Prentice-Hall, Englewood Cliffs, 1961.

行列式, 向量空间, 线性方程组, 线性映射和矩阵, 双线性型和二次型, 特征空间 (没有 Jordan 标准形), 二次曲面的分类, Hilbert 空间的概念和完全连续算子. 特别初等和特别清晰的文笔, 如同 Gel'fand 的和 Halmos 的书, 面向分析和 Hilbert 空间远胜于面向数学的其他分支. 仅有的严重的不足 (此外在 Gel'fand, Mirsky 以及许多我们没有列出的书中也会发现) 是行列式理论对于建立有限维向量空间的性质的应用 (特别是所有的基有同样多的元素这一定理). 但是这点不足如果希望的话容易补救, 而且仅涉及著作的一小部分. 强烈推荐初学者阅读.

[25] B. L. van der Waerden, *Modern Algebra*, 2 vols., Springer, Berlin, 1955.

[26] B. L. van der Waerden, *Modern Algebra*, 2 vols., F. Ungar, New York, 1950.

这部著名的著作首版于 1930 年, 在很长的时间内是 “近世代数” 仅有的完整的陈述. 极其清晰和简洁的风格. 水平比前面列出的各个著作高出许多. 只适合于希望从事数学研究的大学生.

[27] O. Zariski and P. Samuel, *Commutation Algebra*, 2 vols., Van Nostrand, Princeton, 1958.

叙述了 “高等” 算术 (代数数论) 和代数几何的所有必须材料: 域的扩张, Noether 环, Dedekind 整环, 赋值环, 等等. 跟上一部书一样, 掌握了我们的书的内容并且希望成为专家的读者可以懂得本书.

习 题 集

前面列出的一些著作, 特别是 Birkhoff-Maclane, Hoffman-Kunze, Mirsky 的书都包含习题. Hasse 的书伴有一本习题集 (附有答案), 其风格是 “抽象代数” 的, 把它推荐给未来的数学家:

[1] H. Hasse and W. Klobe, *Aufgabensammlung zur Höheren Algebra*, Walter de Gruyter (Sammlung Göschen), Berlin, 1952.

[2] H. Hasse and W. Klobe, *Exercises to Higher Algebra*, F. Ungar, New York, 1954.

当前涉及实际且有效的计算的两本最好的习题集是:

[3] D. K. Fadeev and I. S. Sominskjĭ, *Sbornik zadač po vyssšeĭ algebre*, Fizmatdat, Moscou, 1962.

[4] D. K. Faddeev and I. S. Sominskii, *Problems in Higher Algebra*, W. H. Freeman, San Francisco, 1965.

[5] I. V. Proskurjakov, *Sbornik zadač po vyššeĭ algebre*, Gostekhizdat, Moscou, 1962.

第一本习题集 (980 个题目, 适当处带有详细解答) 几乎涵盖我们的书论述的所有主题; 第二本习题集 (1753 个题目, 适当处带有详细的解答和十分简短的解释) 仅涵盖线性代数, 但是要详尽得多. 我们书的大部分 "数值的" 习题摘自前两部书 (尤其是 [5]).

在法语范围内, 在专门数学教程 (例如在 Lentin 和 Rivaud 的书里) 内自然可以找到许多习题. 另外, 下列针对一般数学的学生的著作包含代数的百余个题目和详尽的解答:

[6] G. Lefort, *Algèbre et Analyse. Exercices*, Dunod, Paris, 1961.

记号索引

对于在正文中引进的记号, 我们指出它第一次出现的节和小节. 对于在习题中定义的记号, 我们指出它所属的节, 并在括号内指出定义该记号的习题编号.

术语索引

对于在正文中引进的术语, 我们指出它第一次出现的节和小节. 对于在习题中定义的术语, 我们指出它所属的节, 并在括号内指出定义该术语的习题编号.

E

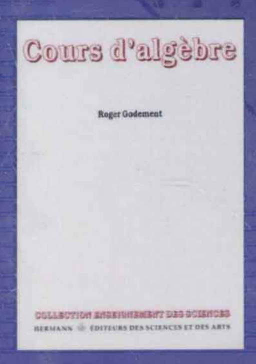
F

K

L

M

Z



本书为法国最好的代数学教科书之一，被誉为代数学教程的“圣经”。

本书以作者在巴黎为大学本科生讲授代数学课程的讲义为基础，内容涵盖了几乎所有本科生需要掌握的，也是未来的数学家和物理学家不可或缺的代数学基础知识：集合和函数、群、环、域、复数；向量空间、线性映射、矩阵；有限维向量空间、线性方程组、行列式、Cramer 公式；多项式、有理分式、代数方程；矩阵的化简等。

本书秉承了法国布尔巴基学派的风格，以专业数学家的语言、现代的观点表述书中的内容，明确严格地定义数学术语，清晰地陈述定理，尽可能完整地证明几乎所有的定理。

本书提供了大量的各种类型的习题，可供不同程度的读者选用， 而且书的最后提供了精心准备的参考文献，帮助读者了解其他观点并养成查询参考书的习惯。

■ 学科类别：数学
academic.hep.com.cn


ISBN 978-7-04-028757-8
9 787040 287578 >
定价 89.00 元